Master of Science Thesis in Electrical Engineering
Department of Electrical Engineering, Linköping University, 2016

Automatic Detection and Classification of Permanent and Non-Permanent Skin Marks

Armand Moulis

LINKÖPING
UNIVERSITY

Master of Science Thesis in Electrical Engineering

Automatic Detection and Classification of Permanent and Non-Permanent Skin Marks

Armand Moulis

LiTH-ISY-EX--YY/NNNN--SE

Supervisor: **Martin Danelljan, ISY, Linköpings universitet**

Examiner: **Lasse Alfredsson, ISY, Linköpings universitet**

Division of Automatic Control
Department of Electrical Engineering
Linköping University
SE-581 83 Linköping, Sweden

Sammanfattning

När forensiker försöker identifiera förövaren av ett brott använder de individuella ansiktsmärken när de jämför den misstänkta med förövaren. Ansiktsmärken används ofta vid identifikation och de hittas idag manuellt. För att skynda på denna process, är det önskvärt att detektera ansiktsmärken automatiskt. Detta examensarbete beskriver en metod för att automatisk detektera och separera permanenta och icke-permanenta märken. Den använder en snabb radial symmetri algoritm som ett huvud element i detektorn och en stödvektormaskin för märkes klassificeraren. Resultatet visar att ansiktsmäkedetektorn har en känslighet på 64% men endast en precision på 8%. Klassifiseraren å andra sidan har en träffsäkerhet på 90% med relativt få åtskilljande kännetecken.

Abstract

When forensic examiners try to identify the perpetrator of a felony, they use individual facial marks when comparing the suspect with the perpetrator. Facial marks are often used for identification and they are nowadays found manually. To speed up this process, it is desired to detect interesting facial marks automatically. This master thesis describes a method to automatically detect and separate permanent and non-permanent marks. It uses a fast radial symmetry algorithm as a core element in the mark detector and a support vector machine for the classifier of the marks. The results shows that the facial mark detector has a 64% recall but only a precision at 8%. The classifier on the other hand shows a 90% accuracy with relatively few discriminative features.

Acknowledgments

I would like to thank the supervisors at NFC in Linköping for their support and guidance during this master thesis. I also want to thank Lasse Alfredsson and Martin Danelljan at Linköpings university for their help during the master thesis process. A special thanks goes to my fiancée for her unconditional love and support through my work.

Linköping, January 20, 2017
Armand Moulis

Contents

Notation

Abbreviations

Abbreviation	Meaning
NFC	National Forensic Centre
RPPVSM	Relatively Permanent Pigmented or Vascular Skin Marks
HOG	Histogram of Oriented Gradients
LBP	Local Binary Patterns
LoG	Laplacian of Gaussian
RGB	Red Green Blue
HSV	Hue Saturation Value

1 Introduction

1.1 Background

The work of systematically recording physical measurements for law enforcement was introduced by Alphonse Bertillon as early as in the 19th century. He developed the Bertillonage system since he believed that each person could be uniquely identified by a set of measurements [1]. This system was however outdated quickly thanks to the explosion of technology.

Today, the amount of video surveillance cameras, security cameras and cellphone cameras increases rapidly and there exist millions of devises capable of catching perpetrators in the act. The videos and still images can be used as evidence for identification during trials where forensic experts evaluate the strength of evidence whether if the suspect is the same person as the one caught on camera.

One common method of evaluating whether the perpetrator and the suspect are the same person is to compare facial features such as eyes, nose, mouth, scars, and other facial marks. This is nowadays done manually [32] by the forensic examiners, and in order to evaluate the strength of the results, a likelihood ratio [21] from Bayes rule is calculated. The likelihood ratio is estimated from two hypotheses, where the numerator gives the probability to achieve the results if the perpetrator and the suspect are the same person. The denominator gives the probability to achieve the results if the perpetrator is another man.

National Forensic Centre (NFC) is currently running a project where an automatic facial recognition system can be used to extract statistics from a database of facial images. The main advantages of using such a method are that the likelihood ratio can be calculated based on statistics, and that the risk for human bias in the decision process is diminished.

Facial features are divided into two groups: class and individual characteristics [29]. The class characteristics includes traits which put individuals into larger groups. Some of these feature are e.g. hair and eye colour, overall facial shape and size of the ears. The class characteristics does not suffice to identify unique individuals. Individual characteristics includes are traits that are unique to an individual, for example the number and location of facial skin marks.

Facial skin marks are any salient skin region that appears on the face. The most common facial marks are moles, pockmarks, freckles, scars, and acne. Some of these marks are not permanent, e.g. acne usually heals without leaving any permanent marks, while scares and moles remain the whole life [20]. Skin marks which can be used for identification need to be relatively permanent, common and also be observable without any special imaging or medical equipment. These relatively permanent marks usually occur due to increased pigmentation or vascular proliferation. Therefore these kind of facial skin marks are called "relatively permanent pigmented or vascular skin marks (RPPVSM)". [23] This master thesis will separate facial skin marks into two classes: permanent and non-permanent facial marks. Which class a facial skin mark belong to is decided by the forensics forensics at NFC.

This master thesis was started due to the need of combining the automatically calculated likelihood ratio value with the evidential value derived from the frequency of facial marks in certain regions of the face. The NFC in Sweden is supporting this work by providing guidance and practical help.

1.2 Motivation

A research relevant to this master thesis is the work by Vorder Bruegge et al. [20] which proposed a fully automatic multiscale facial mark system. It detects facial marks which are stable across the the RGB-channels and different scales. These scales are called Gaussian pyramid and consist of low-pass filtered and subsampled images of the original image. This method to detect permanent marks are also used by Nisha Srinivas et al. [30] who tries to separate identical twins with an automatic multiscale facial mark detector. This method does not try to separate permanent and non-permanent facial mark rather tries to detect the more permanent marks.

An other option considered when looking for facial marks are object detection and object classification. The research on object detection and object classification is a wide and relevant field. Some of the things researchers have tried to detect and classify are faces [2], pedestrians [7] and vehicles [9]. These examples uses descriptive features based on histogram of oriented gradients (HOG) and local binary patterns (LBP). Face detectors also uses Haar-like features [34]. These three sets of features all describes the shape and structure of the searched object.

Taeg Sang Cho et. al. [6] proposed a method using a Support Vector Machine (SVM) as a classifier to separate true and false mole candidates. They used a gist-descriptor as descriptive features. The gist-descriptor is designed to describe texture patterns over space. Read more about the gist-descriptor in the work of Antonio Torralba et. al. [33].

An other work using classifier are the work from Arfika Nurhudatiana et al. [22]. They tried to detect and separate RPPVSM from non-RPPVSM on back torsos. They tried out three different classifiers which include a SVM, neural network and a binary decision tree. As input, the classifier was give the same set of features which included contrast, shape, size, texture, and color. Tim K. Lee et. al. [17] also used the some kind of features but does not use a trained classifier to separate true and false moles on back torsos. They use unsupervised algorithm to classify the mole candidates.

The work [6, 22, 17] all try to separate skin marks and they use a fixed set of features to do this. Arfika Nurhudatiana et al. compare different classifiers but there have been little work on comparing different set of features to separate permanent and non-permanent skin marks. This is why this thesis work will focus on comparing different features as input to a supervised classier. Since the facial marks have a circular shape and mostly vary in color it would bee wise to use colors maps as features.

When it comes to the detection of potential skin mark, there often involves some kind thresholding of a edge enhanced images. Using Laplacian of Gaussian (LoG) kernel as edge enhancement is popular method [10, 24]. After the edge enhancement of a image, the skin marks are highlighted and can then be segmented with different thresholding methods.

This master thesis will however look at a recently used [30, 20] and interesting method to highlight the skin marks. The algorithm is called fast radial symmetry (FRS) and it highlights radially symmetrical regions and suppresses regions that are asymmetrical. This is ideal when one is looking for circular objects which is perfect since facial marks are often circular.

1.3 Aim

The aim of thie master thesis is to create a algorithm that can help create a large data base with facial images and the location of facial marks. By using this algorithm, the evidential value in forensic facial image comparison examinations can be based better. The algorithm should detect facial marks from a color image and then separate them into a permanent and non-permanent group. An overview of the algorithm is presented in figure 1.1. The algorithm consist of several submodules which has their own functionality.

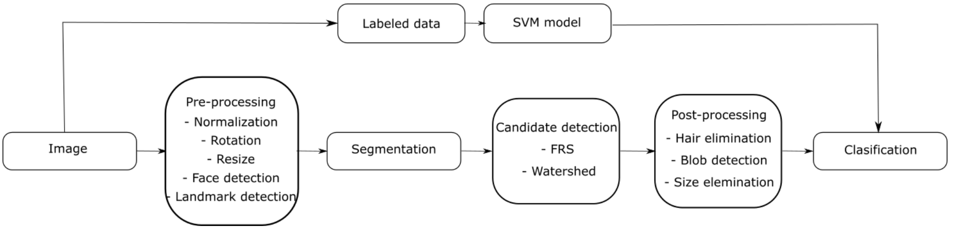

Figure 1.1: *Overview of the algorithm*

1.4 Problem specification

From a single RGB-image of a face en, facial marks should be detected and classified as a permanent or non-permanent mark. This task can be divided into four smaller tasks. These task will be described more in detail later in the master thesis.

Task 1: Pre-processing The image can be illuminated unevenly and rotated which can cause difficulties when detection potential facial marks. Thus, the image has to be geometrically and photometrically normalized.

Task 2: Candidate detection The actual detection of potential marks are done with the help of radial symmetry in the image. The algorithm will search for areas which contains edges that have a cylindrical shape.

Task 3: Post-processing Among the potential facial marks, there can be many false detections such as nostrils, facial hair, pupils et cetera. The false detection has to be eliminated and will be done with a hair removal method, blob identifier, size eliminator and face segmentation.

Task 4: Classification When the marks has been detected, they have to be sorted into the two classes, permanent and non-permanent. This is done by calculating different descriptive features. These features is used to train a supervised support vector machine. With the trained classifier, the facial mark can sorted.

1.5 Scope

In general, when working with image, the quality of the images are crucial for the results. Low resolution and badly illuminated images taken from different angles can cause analytical difficulties. Therefore, this thesis assumes images which are high resolute, well illuminated, taken en face and in RGB-colours.

Also, this master thesis will focus on a comparison between different sets of features for the classifier instead of examining different ways of detecting facial marks. This is due to the little work done regarding feature selection.

The classifier will only be a binary classifier because no non-facial marks has been collected as labelled data during the thesis work due to lack of resources.

1.6 Thesis outline

This chapter describes the aim and problem specification of this master thesis. In Chapter 2, gives an insight in theory behind the methods used in the algorithm. Chapter 3 describes the pipeline of the algorithm and the implementation of the theory used in it. The results from the algorithm can be studied in Chapter 4 and an discussion about the result and methods used is found in Chapter 5. Finally, Chapter 6 consist of a conclusion of the master thesis and ideas for future work within the same scope.

2

Theory

3
Method

This thesis report present a algorithm to detect facial marks automatically. This chapter will briefly describe the method used to develop the algorithm. Further details can be found in the referred works.

3.1 Detection

....

3.1.1 Face detection

An important component for further processing is the bounding box of the face in each image. It is found by using an OpenCV [5] implementation of object detection by Paul Viola et al. [34]. This face detection algorithm was chosen since it has equivalent positive result as other methods [28, 31]. In addition, it is much faster the the other detector. The algorithm from Paul Viola et. al. take advantage of three different parts.

The first is part is a new image representation which allows Haar features from each image to be calculated rapidly. The speed is achieved by using integral images instead of the original image.

The second part is the extraction of the most important features through AdaBoosting. It creates a strong classifier by combining the strength of weaker classifiers. A weak classifier is the best threshold for a feature which separates the faces from the non-faces.

The third part is a cascade decision which reduces the computations costs by rejecting potential bounding boxes for the face. A simple classifier is used to determine if the bounding boxes are promising candidates before a more complex classifier is engaged. This is repeated until all classifiers has been passed or if one of the returns a negative result. All bounding boxes which has returned a negative result is rejected immediately.

When a suitable bonding box as been found for I, the landmarks in the image can be extracted.

3.1.2 Landmark detection

To process a facial image, it is useful to know where different parts of the face are located, e.g. mouth and eyes. These parts can be pinpointed with points called landmarks. With these landmarks, it is possible to create a unique mask for each face and produce a grid with different regions in the face. The landmarks are extracted by using an implementation based on Vahid Kazemi et al. [14]. It uses sate of the art algorithms for face alignment where cascade of regression functions are crucial for its success. The estimated shape of the face is updated by regressing the shape parameters based on normalized features from the image. The parameters are updated until they converges.

From this algorithm, 68 landmarks are extracted where the eyes, mouth, nose and chin are marked. From these, a mask generated where the nostrils, eyes, throat and background are cut out.

3.1.3 Facial grid

The landmarks are also used to produce a grid over the face. The grid consist of 16 regions which are defined the supervisors at NFC. This grid is needed to calculate the number of facial marks within these predefined regions. This is necessary to improve the evidential value of the likelihood ratio.

3.1.4 Normalization

In order to get a reliable and uniform result, the image has to be normalized. There are two kinds of normalization applied on the image, geometric normalization and photometric normalization.

The photometric normalization is performed by using tone mapping operator based on the work of Nikola Banic et al.[3]. It uses a Light Random Sprays Retinex (LRSR) which is an improvement of the Random Sprays Retinex (RSR)[25]. All tone mapping operator transform pixel intensities depending on its surrounding. The RSR uses a random selection of pixels around the the current pixel which

Figure 3.1: *Image over the landmarks (blue points) and facial grid (black lines)*

decreases computations costs, sampling noise and dependency. The calculations are done on the intensity image of each RGB colour channels.

The geometric normalization consist of rotation of the image such that the line between the pupils is aligned with the bottom of the image. The rotations angle is calculated with the help of the landmarks at the corners of the eyes. An angle is calculated by averaging an angle from the right and left eye. The geometric normalization also include a resize of the image such that the interpupillary distance is 500 pixels. A resizing factor is calculated by taking the fraction between 500 and the distance between the pupils.

3.1.5 Segmentation

When searching for facial mark, hair lines and hair can cause false detection. Therefore the image has to be segmented so that only skin area is regarded in search for facial marks. Since interactive segmentation methods are more and more popular [4], it should be beneficial to chose a interactive segmentation method. Carsten Rother et al.[26] compared several popular interactive segmentation methods and also presented their own method, GrabCut. They concluded that GrabCut performs as well as GraphCut [4] with fewer user interactions.

Thus, the segmentation method used for the algorithm is GrabCut which uses

Figure 3.2: *Image after photometric and geometric normalization*

Gaussian Mixture Model (GMM) for a colour image. GrabCut needs a GMM for a know foreground and one for a known background. The known foreground used e.g. is the cheeks and forehead, is extracted with the help of the landmarks.

After creating GMM:s, a energy function is created so that its minimum correspond to a good segmentation which depends on the given foreground and background. The function is minimized iteratively until a converged segmentation is produced.

The segmentation mask is used to improve the mask created from the landmarks. Now using the improved mask, a well segmented image can be searched for facial marks. This image will henceforth be called I_{pre}

3.1.6 Fast Radial Symmetry Detector

There are many ways to extract interesting points or marks. One way is to look at the radial symmetry in the image. This method has been used by several researcher [30, 18, 20, 27]. It seems to be a reliable method since the point is to detect small circular shapes, which is what Jan Schier et. al.[27] did whey they tried to count yeast colonies. This is why the actual mark detector uses an algorithm called Fast Radial Symmetry (FRS) and it was created by Gareth Loy et al.[18].

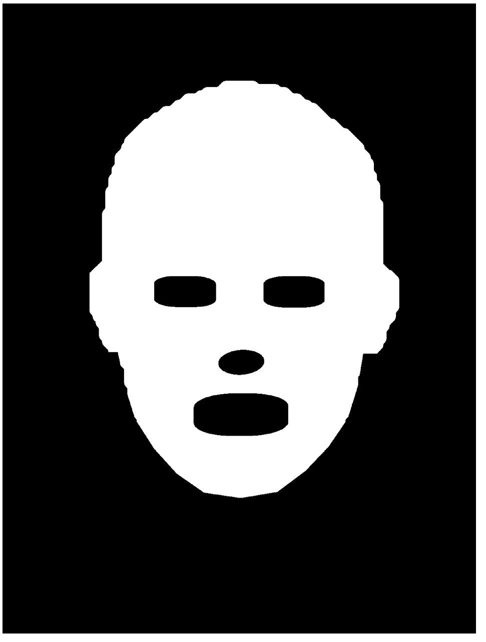

Figure 3.3: *Image of the facial mask*

For each point, p, in the I_{pre}, the contribution of radial symmetry at radius r is calculated by producing an orientation projection image O_n and a magnitude projection image M_n. These images are created by examining the positively-affected, $p_+(p)$, and negatively-affected, $p_-(p)$. To do this, the gradient, g, of the image is needed and it is calculated using a 3x3-Sobel kernel. Since the gradient computations are discrete, it is necessary to average I_{pre} with a 3x3 Gaussian kernel to remove sharp edges.

$$p_+(p) = g(p) + round\frac{g(p)}{\|g(p)\|}n \tag{3.1}$$

$$p_-(p) = g(p) - round\frac{g(p)}{\|g(p)\|}n \tag{3.2}$$

To retrieve the nearest integer the operation $round$ is used. The O_n and M_n are then updated according to eqs. (3.3) to (3.5) and (3.10)

$$O_n(p_+) = O_n(p_+) + 1 \tag{3.3}$$

$$O_n(p_-) = O_n(p_-) - 1 \tag{3.4}$$

$$M_n(p_-) = M_n(p_+) + \|g(p)\| \tag{3.5}$$

$$M_n(p_-) = M_n(p_-) - \|g(p)\| \tag{3.6}$$

The radial symmetry contribution at radius n depends on F_n and A_n which is defined as

$$F_n = \frac{M_n(p)}{k_n}\left(\frac{|\tilde{O}_n(p)|}{k_n}\right)^{\alpha} \tag{3.7}$$

$$\tilde{O}_n(p) = \begin{cases} O_n(p) & \text{if } O_n(p) < k_n \\ 0 & \text{else} \end{cases} \tag{3.8}$$

A_n is a Gaussian kernel with different size depending on n, α is radial strictness parameter and k_n is a scaling factor. α is set to 2 and k_n to 9.9 since Gareth Loy et al. deemed suitable for most applications.

The final radial symmetry image S_n is calculated

$$S_n = F_n * A_n \tag{3.9}$$

This was a calculation for radius n and it desirables to use multiple radii to detect point larger than n. It is not necessary to use a continuous spectrum of radii, thus the radii used are $N = \{1, 3, 5, 7, 9, 11, 13, 15\}$

The average of radial symmetry images, S, are calculated as

$$S = \frac{1}{N}\sum_{n=1}^{N} S_n \tag{3.10}$$

At this point, an FRS-image with points of interest has been acquired. From this image, a binary threshold was applied with the threshold h_{FRS} to extract sinks. The sinks are needed for the watershed algorithm described by Fernand Meyer [19]. The use of watershed is good since it can find the contour of uneven marks as long as the pixels approximately have the same intensity value. The watershed algorithm with the sinks are applied on a grey image of the face. The output from this is a set of bonding boxes containing facial marks. This set is henceforth called candidates and the module used to detect the facial marks is called FRS-detector.

3.1.7 Candidate elimination

Since many facial mark candidates may be false positives, they have to be discovered and excluded. Vorder Bruegge et al. [20] used three elimination methods which seemed intuitive. Size, shape and presence of hair should be good indicators if the candidate is a false detection or not. Each detected candidate is given a 30x30 area which is processed through three eliminators.

Facial marks are often blob-shaped which is why the first eliminator uses a simple blob detector from OpenCV. It creates a thresholded images with connective pixels and does this with different threshold values. If the union of all the different images does not contain a blob-shaped object, the area is excluded from the candidates.

The second eliminator uses a hair removal algorithm by Tim Lee et. al. [16]. The algorithm smooth out hair pixels with closing operations using the three different structuring elements. The suggested structuring elements by Tim Lee et. al. is larger than the one used in this implementation since their hair-structures where wider. Thus, the smaller structuring elements T_0, T_{45} and T_{90} were used.

$$T_0 = \begin{bmatrix} 0 & 1 & 1 & 1 & 0 \end{bmatrix} \quad T_{45} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad T_{90} = (T_0)^T$$

The closed image is generated by applying each structure element on each colour channel as (3.11), where G is the closed image, M is the image of a mark, $T_x = [T_0, T_{45}, T_{90}]$ and C the RGB-channels. This means that M_c is a gray image of a mark where the structuring elements detect thin and small edges.

$$G_c = |M_c - \max_x(M_c * T_x)| \tag{3.11}$$

$\max_x(M_c * T_x)$ means that the largest pixel value from the structuring elements are pick for that colour channel. If the number of hair pixels in the union of G_c is larger then a threshold h_{hair}, the mark is excluded from the candidates.

The third eliminator removes candidates depending on their size. If the candidate has an area smaller than 5 pixels or an area larger than 1000 pixels, it is eliminated. The thresholds were chosen because since all annotated marks is within this interval, see fig. 3.4.

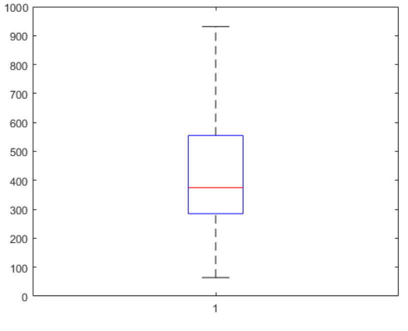

Figure 3.4: *The distribution of areas from the annotated facial marks.*

3.2 Classification

When choosing the suitable classifying method for the mark classifier, there are many methods to chose from. Kotsiantis et. al. [15] concludes that the best machine learning method depends greatly on the conditions the classifier are going to work in. Generally, support vector machine (SVM) tend to perform better with continuous and multidimensional features and with a large amount of samples. The features used in this case fulfill the continuity and multi dimension. Chih-Wei et. al. [11] also describes a good way to optimize the use of SVM and recommend to use radial basis function (RBF) as kernel. This is why a SVM with RBF kernel is chosen for the mark classifier.

The goal with a SVM is to separate different classes by finding the best hyperplane which divides them. The hyperplane is moved such that a loss-function is minimized. RBF kernel nonlinearly maps samples into a higher dimensional space which allows the classifier to handle non-linearly separable classes. This kernel also has fewer numerical difficulties. The parameters needed for a RBF kernel i C which determines the penalty parameter for the error and γ which defines how far the influence of a single training sample reaches. [11]

The training data consists of the labelled facial marks provided by the supervisors at NFC. To get a good classifier, a set of discriminative features are required.

3.2.1 Features

The most common color space in use is the RGB system, one channel each for the red, green and blue colors. Arfika Nurhudatiana et al. [22] used, among others, the minimum, maximum, and average from the RGB-channels as discriminative features. By using 11 more colour channels, presented by Joost van de Weijer et al.[13], it is possible to improve the classifier. The colour transformation is a trained colour mapping. It is trained on real world images from Google Image and has shown to out perform colour chips. The 11 colours consist of black, blue, brown, grey, green, orange, pink, purpale, red, white and yellow.

The features extracted from the facial marks are the mean and the standard deviation from the three RGB channels and the 11 colours from the work of Joost van de Weijer et al. This results in 28 features which is used to train the classifier.

To not let some feature with greater numeric rang dominate over features with smaller range, the features need to be scaled [11]. This is very important which is why the features are linearly scaled to a range from 0 to 1. The same scaling factor has to be used when the test data is scaled.

4

Result

This Chapter first describes the experiment to evaluate the algorithm and then presents the results.

4.1 Experiment

To evaluate the algorithm, a set of 106 images of faces en face were acquired from SCface database [8]. Facial marks of interest were marked and labeled as a permanent or a non-permanent by the supervisors at NFC. This resulted in 506 marks where 353 were permanent and 153 were non-permanent.

The experiment was set such that the image set was processed by the algorithm with 11 different thresholds values, h_{FRS}, for the FRS-image. The h_{FRS} ranged from 0.05 to 0.15. The output was compared to the ground truth. A correct detection was defined as all detections which had a union with an annotated mark. This definition has been chosen since some of the detections can be very small. Also, since candidates larger than 1000 pixels has been eliminated, no over large candidates can give correct detections.

The evaluation measurement for the detector is the precision (4.1) and recall (4.2) value. Precision tells how well the detector is to avoid false detections while recall tells how well it finds the annotated marks. The result from the different h_{FRS}-values can be seen in fig. 4.1.

$$Precision = \frac{\text{Number of correct detections}}{\text{Number of detections}} \tag{4.1}$$

$$Recall = \frac{\text{Number of correct detections}}{\text{Number of annotated marks}} \tag{4.2}$$

The h_{FRS}-value which gives the best recall value was used to evaluate the elimination process of the candidates. This was done by calculating the precision and recall values before the different elimination steps. The results is displayed in fig. 4.1.

To evaluate the facial mark classifier, a cross validation of the 506 annotated mark were performed. 25 marks was chosen at random to be used as test marks while the remaining marks was used for training the SVM. This was repeated until all the marks had been used as test marks.

In order to find the best C-value and γ-value for the mark classifier, the parameters are varied over a rough interval to narrow down the search. Afterwards, a more fine interval is used to find the best parameters. This has been shown by Chih-Wei et. al. [11] to be an effective method compared to a more random selection of parameters which is often used by people unfamiliar to SVM.

4.2 Results from experiment

In the fig. 4.1, the precision and recall for different h_{FRS}-value can be examined. The precision corresponds to the white bar and the recall corresponds to the black bar. Note that this is only the detections of facial mark and no classification between permanent and non-permanent marks.

As one can see, the precision increases with higher h_{FRS}-value without affecting the recall substantially. This is means that the number of candidates decreases with a growing h_{FRS}-value. Thus, a small h_{FRS}-value results in a large amount of candidates while a larger value gives fewer candidates.

In fig. 4.2, it is possible to see the effects of the different elimination steps. As before, the white bar represent the precision and the black bar represent the recall. The first pair is the result just after the candidate detection and the second pair is the result after the blob detector. Furthermore, the third pair is after the hair eliminator and the last pair is after the size eliminator.

It is obvious that the different eliminators are essential for the algorithm. The hair eliminator improves the precision the most while the blob detector worsen the recall the most.

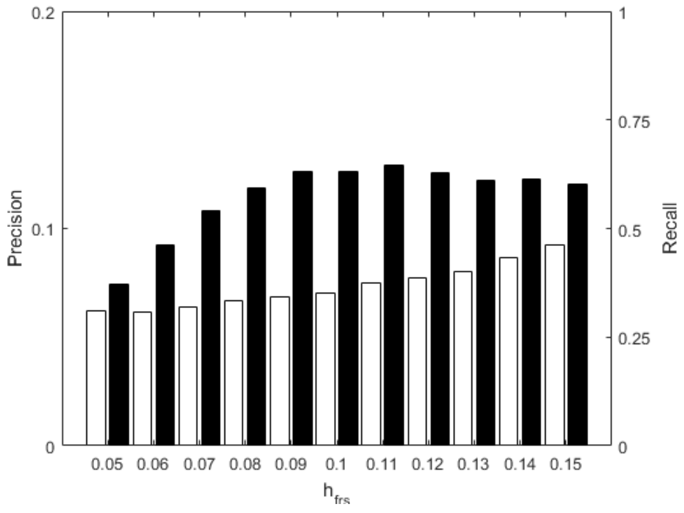

Figure 4.1: *Detection results from the algorithm with different h_{frs}-values. The white bars represent the precision value and the black bars represent recall value.*

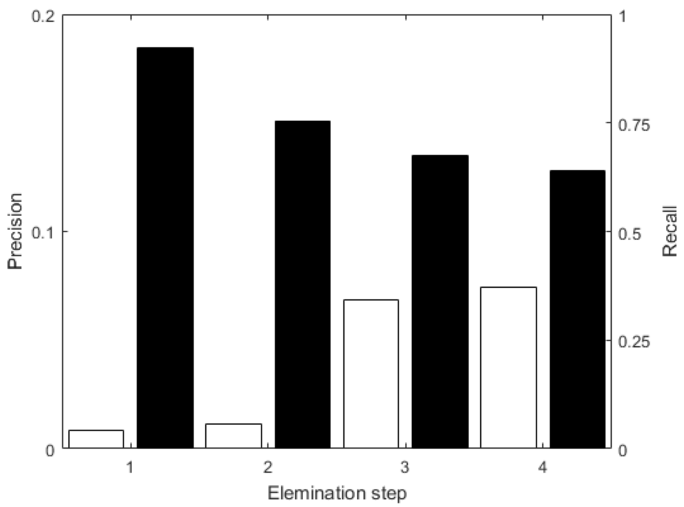

Figure 4.2: *Detection results from the algorithm after different candidate elimination steps. 1 = before elimination, 2 = after blob-elimination, 3 = after hair-elimination, 4 = after size-elimination. The white bars represent the precision value and the black bars represent recall value.*

The result from the mark classifier is presented as accuracy matrix with varying

C-value and γ-value. The accuracy is calculated as in (4.3). In fig. 4.3, the C-value ranges from 1 to 100 while the γ-value ranges from 10^{-2} to 1.

$$Accuracy[\%] = \frac{\text{True positive} + \text{True negative}}{\text{Number of annotated marks}} \tag{4.3}$$

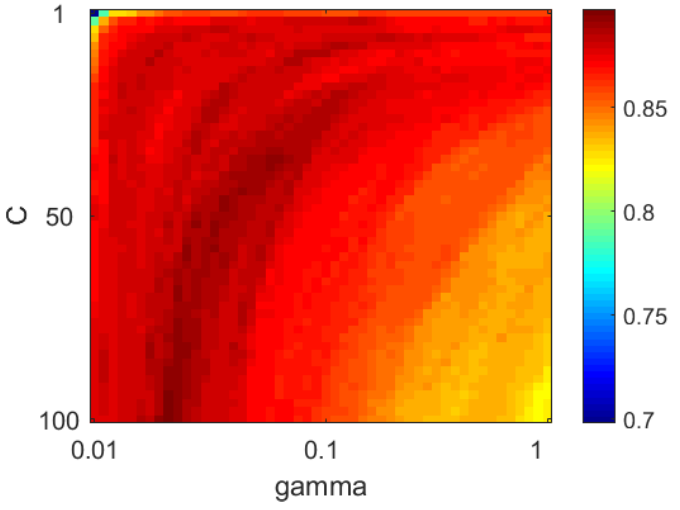

Figure 4.3: *Crude accuracy matrix where* $1 \leq \gamma \leq 100$ *and* $10^{-2} \leq \gamma \leq 1$.

As show, the best accuracy is when $C \approx 95$ and $\gamma \approx 5 * 10^{-2}$. Therefore, with a finer interval for the two parameters, fig. 4.4 was generated. Here, the C-value ranges from 90 to 115 while the γ-value ranges from $5 * 10^{-2}$ to $2 * 10^{-1}$.

Form fig. 4.4 it is possible to conclude that the best accuracy is acquired with several pair of parameters and results in the accuracy 90%. $C = 100$ and $\gamma = 0,14$ is one of those pairs. Therefore, the mark classifier is given these values as parameters for the SVM.

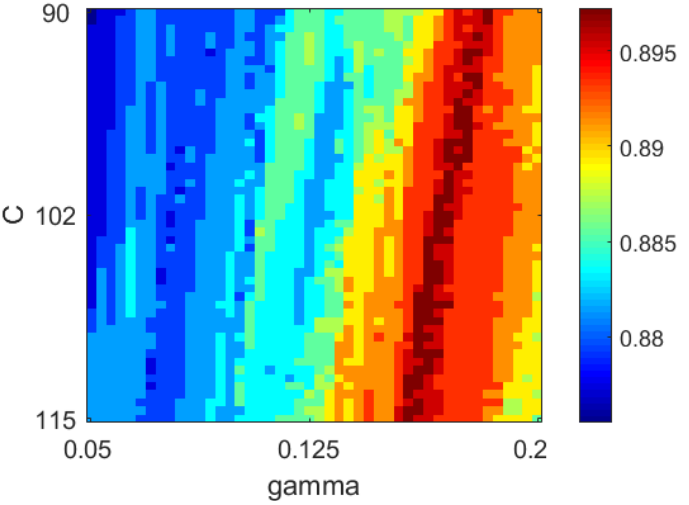

Figure 4.4: *Fine accuracy matrix where* $90 \leq \gamma \leq 115$ *and* $5 * 10^{-2} \leq \gamma \leq 2 * 10^{-1}$.

5 Discussion

This section will discuss the results from the algorithm and the methods used to implement it.

5.1 Result

As seen, the detector has problem with false detections which is a huge problem. The precision is not very good, not even over 10%, due to the many false detections. The precision does not increase faster than the decline of the recall with an increasing h_{frs}-value. This indicates that there are margins for improvement when it comes to the candidate detector. The elimination method used does improve the precession well enough which means that they also can be improved.

When looking at the accuracy of the mark classifier, with an accuracy of 90%, the result is pretty satisfactory.

5.2 Method

There are a lot to say about the methods used in the algorithm. The major problem with the algorithm is the elimination of candidates. The blob detector works well in not eliminating true detections which the hair removal algorithm does not. It eliminates candidates which are true facial marks. This is because it indicates that the facial marks are hair which makes it hard to separate the true candidates and the hair intensive candidates. Tim et Lee al. describes their algorithm well except when they are explaining how to calculate the hair mask for each colour channel. It is not clear what the maximum from refers to. The algorithm in this work used the maximal pixel value between the different structuring elements.

The mark detector used in the algorithm was good at indicating the potential facial marks but the simple thresholding method to pin point them out was not optimal. It kept the pixels larger than a certain percent of the maximal value in the FRS image. This resulted in many unnecessary candidates which of course contributed to the high false detection rate.

The blob-detector used is hardly improving the precision at the cost of recall loss. This means that the blob-detector is not contributing to the algorithm in a positive way. The hair-eliminator on the other hand does improve the precision which is the whole point with the candidate eliminators.

When it comes to the mark classifier, it could have performed better if the number of features was larger and more explored. Now the features are very simple and there is overweight of permanent marks among the annotated facial marks which is not ideal.

Regarding the references used in these thesis, several of them uses FRS to detect point of interest which shows the actuality of the method. Many of the papers trying to detect facial marks uses a crude segmentation mask which does not follow the hairline and chin well. This algorithm uses a more precise segmentation method which reduces the areas which is not processed.

5.3 Ethical perspective

As with all applications which can be used for surveillance of people, the integrity is at stake. Facial recognition algorithms using facial marks can be misused for malicious intent. They can also help the legal system to catch and convict criminals which is desirable outcome of this paper.

When it comes to the facial images, they are taken from a open source database which should only be used for academical research. There are no personal information attached to the images which makes them as anonymous as possible without corrupting the images.

6 Conclusion

This master thesis has examined the possibilities to develop an algorithm which could detect facial marks and separate them into permanent and non-permanent. The result from the experiments shows that the detector can initially find the facial mark with high recall but with a low precision. The precision increases as the false detections are eliminated. The mark classifier demonstrates good result with its accuracy of 90%.

A proposed remedy for the low precision on the detector is to improve the elimination of the false candidates. The hair-detector is to crude and may be combined with or replaced by a module which looks at the Fourier Transform of the candidates. The mark classifier could perform better with better discriminative features. By finding better features through examination of the differences between permanent and non-permanent marks.

Bibliography

[1] Bertillon A. Forensic facial analysis. identification anthropometrique: instructions signaletiques. France, Paris, 1885. Not cited.

[2] Timo Ahonen, Abdenour Hadid, and Matti Pietikainen. Face description with local binary patterns: Application to face recognition. *IEEE transactions on pattern analysis and machine intelligence*, 28(12):2037–2041, 2006. Not cited.

[3] Nikola Banić and Sven Lončarić. Color badger: a novel retinex-based local tone mapping operator. In *International Conference on Image and Signal Processing*, pages 400–408. Springer, 2014. Not cited.

[4] Yuri Y Boykov and M-P Jolly. Interactive graph cuts for optimal boundary & region segmentation of objects in nd images. In *Computer Vision, 2001. ICCV 2001. Proceedings. Eighth IEEE International Conference on*, volume 1, pages 105–112. IEEE, 2001. Not cited.

[5] Gary Bradski et al. The opencv library. *Doctor Dobbs Journal*, 25(11):120–126, 2000. Not cited.

[6] Taeg Sang Cho, William T Freeman, and Hensin Tsao. A reliable skin mole localization scheme. In *2007 IEEE 11th International Conference on Computer Vision*, pages 1–8. IEEE, 2007. Not cited.

[7] Markus Enzweiler and Dariu M Gavrila. A multilevel mixture-of-experts framework for pedestrian classification. *IEEE Transactions on Image Processing*, 20(10):2967–2979, 2011. Not cited.

[8] Mislav Grgic, Kresimir Delac, and Sonja Grgic. Scface–surveillance cameras face database. *Multimedia tools and applications*, 51(3):863–879, 2011. Not cited.

[9] Feng Han, Ying Shan, Ryan Cekander, Harpreet S Sawhney, and Rakesh Kumar. A two-stage approach to people and vehicle detection with hog-based svm. In *Performance Metrics for Intelligent Systems 2006 Workshop*, pages 133–140. Citeseer, 2006. Not cited.

[10] Brian Heflin, Walter Scheirer, and Terrance E Boult. Detecting and classifying scars, marks, and tattoos found in the wild. In *Biometrics: Theory, Applications and Systems (BTAS), 2012 IEEE Fifth International Conference on*, pages 31–38. IEEE, 2012. Not cited.

[11] Chih-Wei Hsu, Chih-Chung Chang, Chih-Jen Lin, et al. A practical guide to support vector classification. 2003. Not cited.

[12] Anil K Jain and Unsang Park. Facial marks: Soft biometric for face recognition. In *Image Processing (ICIP), 2009 16th IEEE International Conference on*, pages 37–40. IEEE, 2009. Not cited.

[13] Jakob Verbeek Diane Larlus Joost van de Weijer, Cordelia Schmid. Learning color names for real-world applications. *IEEE Transactions on Image Processing*, 18(7):1512–1523, July 2009. Not cited.

[14] Vahid Kazemi and Josephine Sullivan. One millisecond face alignment with an ensemble of regression trees. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 1867–1874, 2014. Not cited.

[15] Sotiris B Kotsiantis, I Zaharakis, and P Pintelas. Supervised machine learning: A review of classification techniques, 2007. Not cited.

[16] Tim Lee, Vincent Ng, Richard Gallagher, Andrew Coldman, and David McLean. Dullrazor®: A software approach to hair removal from images. *Computers in biology and medicine*, 27(6):533–543, 1997. Not cited.

[17] Tim K Lee, M Stella Atkins, Michael A King, Savio Lau, and David I McLean. Counting moles automatically from back images. *IEEE Transactions on Biomedical Engineering*, 52(11):1966–1969, 2005. Not cited.

[18] Gareth Loy and Alexander Zelinsky. Fast radial symmetry for detecting points of interest. *IEEE Transactions on pattern analysis and machine intelligence*, 25(8):959–973, 2003. Not cited.

[19] Fernand Meyer. Color image segmentation. In *Image Processing and its Applications, 1992., International Conference on*, pages 303–306. IET, 1992. Not cited.

[20] Richard W. Vorder Bruegge Ph.D. Nisha Srinivas M.Sc., Patrick J. Flynn Ph.D. Human identification using automatic and semi-automatically detected facial marks. *Journal of Forensic Sciences*, 61(S1):117–130, September 2015. Not cited.

[21] Anders Nordgaard, Ricky Ansell, Weine Drotz, and Lars Jaeger. Scale of conclusions for the value of evidence. *Law, probability and risk*, pages 1–24, 2011. Not cited.

[22] Arfika Nurhudatiana, Adams Wai-Kin Kong, Lisa Altieri, and Noah Craft. Automated identification of relatively permanent pigmented or vascular skin marks (rppvsm). In *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 2984–2988. IEEE, 2013. Not cited.

[23] Siu-Yeung Cho Craft N. Nurhudatiana A., Matinpour K. Fundamental statistics of relatively permanent pigmented or vascular skin marks for criminal and victim identification. In *Biometrics (IJCB)s*. Not cited.

[24] Unsang Park and Anil K Jain. Face matching and retrieval using soft biometrics. *IEEE Transactions on Information Forensics and Security*, 5(3):406–415, 2010. Not cited.

[25] Edoardo Provenzi, Massimo Fierro, Alessandro Rizzi, Luca De Carli, Davide Gadia, and Daniele Marini. Random spray retinex: A new retinex implementation to investigate the local properties of the model. *IEEE Transactions on Image Processing*, 16(1):162–171, 2007. Not cited.

[26] Carsten Rother, Vladimir Kolmogorov, and Andrew Blake. Grabcut: Interactive foreground extraction using iterated graph cuts. In *ACM transactions on graphics (TOG)*, volume 23, pages 309–314. ACM, 2004. Not cited.

[27] Jan Schier and Bohumil Kovár. Automated counting of yeast colonies using the fast radial transform algorithm. In *Bioinformatics*, pages 22–27, 2011. Not cited.

[28] Henry Schneiderman and Takeo Kanade. A statistical method for 3d object detection applied to faces and cars. In *Computer Vision and Pattern Recognition, 2000. Proceedings. IEEE Conference on*, volume 1, pages 746–751. IEEE, 2000. Not cited.

[29] Nicole A Spaun and Richard W Vorder Bruegge. Forensic identification of people from images and video. In *Biometrics: Theory, Applications and Systems, 2008. BTAS 2008. 2nd IEEE International Conference on*, pages 1–4. IEEE, 2008. Not cited.

[30] Nisha Srinivas, Gaurav Aggarwal, Patrick J Flynn, and Richard W Vorder Bruegge. Analysis of facial marks to distinguish between identical twins. *Information Forensics and Security, IEEE Transactions on*, 7(5):1536–1550, 2012. Not cited.

[31] K-K Sung and Tomaso Poggio. Example-based learning for view-based human face detection. *IEEE Transactions on pattern analysis and machine intelligence*, 20(1):39–51, 1998. Not cited.

[32] Pedro Tome, Ruben Vera-Rodriguez, Julian Fierrez, and Javier Ortega-Garcia. Facial soft biometric features for forensic face recognition. *Forensic science international*, 257:271–284, 2015. Not cited.

[33] Antonio Torralba, Aude Oliva, Monica S Castelhano, and John M Henderson. Contextual guidance of eye movements and attention in real-world scenes: the role of global features in object search. *Psychological review*, 113(4):766, 2006. Not cited.

[34] Paul Viola and Michael Jones. Rapid object detection using a boosted cascade of simple features. In *Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on*, volume 1, pages I–511. IEEE, 2001. Not cited.