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

Automatic Detection and Localization of Relatively Permanent Pigmented or Vascular Skin Marks

Armand Moulis

LINKÖPING UNIVERSITY

Master of Science Thesis in Electrical Engineering

Automatic Detection and Localization of Relatively Permanent Pigmented or Vascular Skin Marks

Armand Moulis

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

Supervisor: **Doktorand, 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

Abstract

Acknowledgments

Linköping, October 16, 2016
Armand Moulis

Contents

Notation

1

Introduction

1.1 Motivation

The amount of technical tools available for forensic analysis in law enforcement increases rapidly and today there exist millions of devises capable of taking colour images. Video surveillance cameras, security cameras and cellphone cameras can all be used to catch perpetrators in the act. The videos and still images can be used as evidence for identification during trails which means that forensic technicians need tools to evaluate 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 [17] by the forensic examiners and in order to give a objective and comparable conclusion value a likelihood ratio [11] is calculated. The likelihood ratio expresses how strong the evidences against the suspect.

To calculate the likelihood ratio it is required to have enough observations of facial features and these are acquired manually by experts since facial recognition processes have not been found to be reliable enough [10]. To record all these observations manually is time consuming and there exist a interest in doing this automatically [15].

This master thesis was motivated by the need of large amount of data from facial marks. The National Forensic Centre (NFC) in Sweden is supporting this work by providing guidance and practical help.

1.2 Aim

The aim of this master thesis is to examine the possibilities of automatically detecting and locating facial marks and classifying them as permanent or non-permanent marks. The frequency, location and size of the permanent marks are stored such that it can be used to calculate the likelihood ratio. By automatically creating a large data base with face images and their features, the accuracy and speed in face recognition cases can be increased.

1.3 Problem specification

This master thesis is going to answers the following questions:

> Is it possible to implement a algorithm which can automatically detect RPPVSM?
>
> How can the RPPVSM be given a location and size within a face?
>
> With which accuracy can the algorithm detect and localize RPPVSM?

1.4 Boundary

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 will us images which are high resolute, well illuminated, taken en face and in RGB-colours.

2

Related work/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.

Resent research by Srinivas et al. [10] have resulted in an automatic and semi-automatic facial recognition processes. It uses a multiscale automatic facial mark detector for the automatic detector and receive a equal error rate of 15.48%. This result was improved by introducing human knowledge in the semi-automatic detector.

When distinguishing identical twins it is useful to look at facial marks which has been examined in an other article by Srinivas et al. [16]. The study concluded that the facial marks can be uses as features for distinguishing between identical twins even if there seems to exist a correlations between the twins set of marks.

Nurhudatiana et al. [12] describes in their article the distribution of Relatively Permanent Pigmented or Vascular Skin Marks (RPPVSM) in Caucasians, Asians, and Latinos. They conclude that if the number of RPPVSM are few they are randomly distributed which can be used for personal identification in law enforcement.

Anil and Park found during their research [4] that facial mark can be used to increase the recall and precision for a state-of-the-art face matcher (FaceVACS). The facial mark detector used the 3x3 LoG-operator as blob detector. Adding these features to the algorithm improved the face matcher from 92.96% to 93.90% on the Facial Recognition Technology (FERET) database and from 91.88% to 93.14% on a Mugshot face database.

2.1 Facial marks

The skin in the face does not have a homogeneous colour and contains regions with different coloured facial marks. 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 marks, while scares and moles remain the whole life [10]. Skin marks which can be used for identification are called RP-PVSM and they have to be relatively permanent, common and also be observable without any special equipment. [12]

3

Method

This paper present a algorithm to detect facial marks automatically. This section will describe the methods briefly used to develop the algorithm. If more details are requested, they can be found in the referred works.

3.1 Detection

An overview of the detections system is presented in figure 3.1. The system consist of several subsystems which has it own functionality. Each image, I, is processed and go through each of these subsystems.

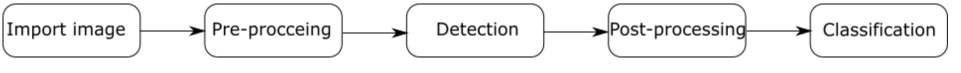

Figure 3.1

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 implementation of object detection by Paul Viola et al. [18]. The algorithm 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. If the first weak classifier gives a positive result, the second classifier is engaged. This is repeated until all weak 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

With landmark in the face, 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 implantation based on Vahid Kazemi et al.[6] 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, moth, nose and chin are marked. From these, a mask generated where the nostrils, eyes, throat and background is cut out. Also, a grid of 16 regions is generated according to the supervisors at NFC demands.

3.1.3 Normalization

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

The geometric normalization consist of rotation of the image such that the eyes are 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.

The photometric normalization is performed by using tone mapping operator based on the work of Nikola Banic et al.[2]. It uses a Light Random Sprays Retinex (LRSR) which is an improvement of the Random Sprays Retinex (RSR)[13]. All tone mapping operator transform pixel intensities depending on its surrounding and the RSR uses a random selection of pixels around the the current pixel which decreases computations costs, sampling noise and dependency. The calculations are done on intensity image for LRSR while uses the RGB colour channels.

3.1.4 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. The segmentation is done with GrabCut produced by Carsten Rother et al.[14]. It uses Gaussian Mixture Model (GMM) for a colour image. GrabCut needs a GMM for know foreground and one for 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 searched for facial marks. This image will henceforth be called I_{pre}

3.1.5 Colour space

Facial marks tend to have a brown or red colour and thus would it be of interest to process these colours in the image. This is possible by converting the RGB image to red-channel and brown-channel images respectively. This is done by using a trained transformer from the work of Joost van de Weijer et al.[5]. It is trained on real world images from Google Image and has shown to out perform colour chips. The colour transformed images are used for the watershed algorithm to find the contours of the facial marks and to extract relevant features for the classifier.

3.1.6 Fast Radial Symmetry Detector

The actual mark detection uses local radial symmetry to detect interesting point in the I_{pre}. The algorithm used i called Fast Radial Symmetry (FRS) and was created by Gareth Loy et al.[8]. For each point, p, in the I_{pre}, the contribution of radial symmetry at radius r is calculated by producing a 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.

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

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

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 calculate

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

This was a calculation for radius n and it desirables to use multiple radii to detect point larger then 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, a image with points of interest has been acquired. From this image, a binary threshold was applied with the threshold h_{FRS} to extract sinks for the watershed algorithm described by Fernand Meyer [9]. The output is now a set of bonding boxed containing facial mark candidates. A small h_{FRS}-value results in a large amount of candidates while a larger value gives fewer candidates.

3.1.7 Candidate elimination

Since many facial mark candidates may be false positives, they have to be discovered and excluded. Each candidate is given a 25x25 area which is processed through two eliminators.

The first eliminator uses a simple blob detector. It searched thresholded binary images for connective pixels and returns true is the area contains a blob. If the mark does not contain a blob, it is excluded from the candidates.

The eliminator uses a hair removal algorithm by Tim Lee et. al.[7]. 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 then 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.

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

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.

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

$max(M_c * T_x)$ means that the largest pixel value from the structuring elements is 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.

3.2 Classification

The algorithm should also separate RPPVSM from non-permanent marks. This is done by training a SVM-classifiers using radial basis function kernels and 1000 iteration. The training data consists of the labelled facial marks provided by the supervisors at NFC. To get a good classifier, a set of features are required.

3.2.1 Features

The features extracted from the facial marks are simply the mean and the standard deviation from the three RGB channels and the red and brown channels.

3.3 Experiment

To evaluate the algorithm, a set of 106 images of faces en face was acquired from SCface database [3] and FRGC database CITE. Facial marks of interest were marked and labeled as a RPPVS or a non-RPPVS by the supervisors at NFC. This resulted in 506 marks where 353 were RPPVS and 153 were non-RPPVS.

The experiment was set such that the image set was processed by the algorithm with 11 different h_{frs}-values ranging from 0.03 to 0.13. The output was compared to the ground truth.

To evaluate the RPPVS 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.

The results from the detector and the classifier is presented in chapter 4.

4

Result

From among the 106 test images, a certain amount of true detections were made. The ratio between these and the 506 annotated marks is called detection ratio. In the fig. 4.1, the ratio of true detections corresponds to the white bar and the number of false detections corresponds to the black bar. Note that this is only the detections of facial mark and no classification between RPPVS and non-RPPVS.

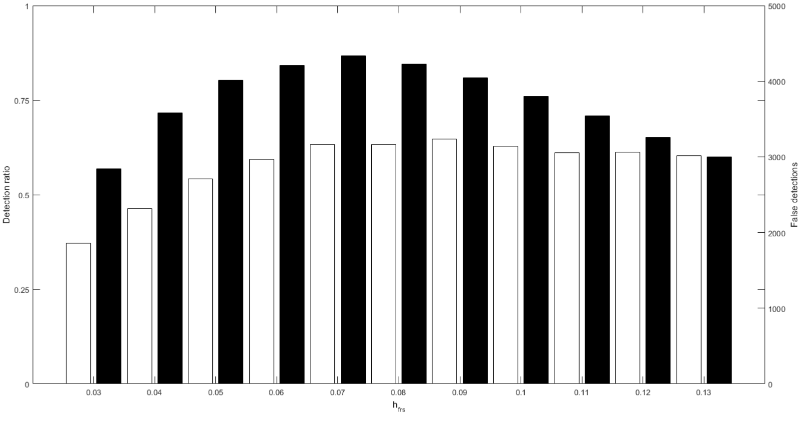

Figure 4.1: *Detection results from the algorithm with different h_{frs}-values. The white bars represent the detection ratio and the black bars represent the number of false detections*

The result from the classifier can be presented as a confusion matrix, see table 4.1. A positive mark is a RPPVS and a negative mark is a non-RPPVS. The accuracy of the classifier is $49,6\%$

	Predicted positive	Predicted negative
Labeled positive	128	225
Labeled negative	30	123

Table 4.1: *Confusion matrix from RPPVS classifier*

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 detection ratio is also not very high, with a maximum of 56%. The ratio does not however decline as fast as the false detections with an increasing h_{frs}-value. This should have been examine more closely if the false detection rate could have dropped even further without affecting the detection ratio.

When looking at the accuracy of the RPPVS classifier, the result is very poor and is not even better than a random classifier. The confusion matrix is however not random but has a tendency to give false positive output.

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 eliminate 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 FRS-detector used in the detector algorithm was good at indicating the potential facial marks but the simple thresholding method to pin point out them was not optimal. It used the 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.

When it comes to the RPPVS classifier, it could have performed better if the number of features was larger and more explored. Now the features are very simple and does not apparently separate the RPPVS and non-RPPVS well. Also, there is overweight of RPPVS in the among the annotated facial marks which could have contributed to the poor classifier.

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 application which can be used for surveillance of people, the integrity is at stake. Facial recognition algorithms using facial marks can to 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

Future work and connection to problem specification

Bibliography

[1] Bertillon A. Forensic facial analysis. identification anthropometrique: instructions signaletiques. France, Paris, 1885. Cited on page 3.

[2] 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. Cited on page 6.

[3] Mislav Grgic, Kresimir Delac, and Sonja Grgic. Scface–surveillance cameras face database. *Multimedia tools and applications*, 51(3):863–879, 2011. Cited on page 9.

[4] 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. Cited on page 3.

[5] 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. Cited on page 7.

[6] 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. Cited on page 6.

[7] 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. Cited on page 8.

[8] 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. Cited on page 7.

[9] Fernand Meyer. Color image segmentation. In *Image Processing and its Applications, 1992., International Conference on*, pages 303–306. IET, 1992. Cited on page 8.

[10] 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. Cited on pages 1, 3, and 4.

[11] 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. Cited on page 1.

[12] 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*. Cited on pages 3 and 4.

[13] 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. Cited on page 6.

[14] 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. Cited on page 7.

[15] Bruegge R.W.V. Spaun N.A. Forensic identification of people from images and video. In *Biometrics: Theory, Applications and Systems*. Cited on page 1.

[16] 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. Cited on page 3.

[17] 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. Cited on page 1.

[18] 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. Cited on page 5.