2013 International Workshop on Biometrics and Forensics (IWBF). IEEE, 2013

DOI: http://dx.doi.org/10.1109/IWBF.2013.6547304

ANALYSIS OF THE VARIABILITY OF FACIAL LANDMARKS IN A FORENSIC SCENARIO

Ruben Vera-Rodriguez[a], *Pedro Tome*[a], *Julian Fierrez*[a],
Nicomedes Expósito[b] *and Francisco Javier Vega*[b]

[a]Biometric Recognition Group - ATVS, Escuela Politecnica Superior
Universidad Autonoma de Madrid
Avda. Francisco Tomas y Valiente, 11 - Campus de Cantoblanco - 28049 Madrid, Spain
{ruben.vera, pedro.tome, julian.fierrez}@uam.es
[b]Dirección General de la Guardia Civil - DGGC Madrid, Spain
{nexposito, jvega}@guardiacivil.es

ABSTRACT

This paper reports an study of the variability of facial landmarks in a forensic scenario. This variability is affected by two factors: on the one hand, the precision in which the landmarks are tagged (manually or automatically), and on the other hand some other variability factors such as the pose, expression, occlusions, etc. For this study, a mugshot database of 50 persons has been collected following the procedure used by the Spanish Guardia Civil. Mugshots are taken with three distances between the persons and the camera (3, 2, 1 meters) showing the full body, the upper body and the face respectively, obtaining in total 1200 images. 21 facial landmarks are defined and the database was manually tagged imitating the procedure followed by a forensic examiner. This paper analyses the facial landmarking variability for the three distances considered, and also considering the differences obtained for male and female. Results show that landmarks located in the outer part of the face (highest end of the head, ears and chin) present a higher level of variability compared to the landmarks located the inner face (eye region, and nose). Regarding the gender, the landmarks placed in the outer part of the face present a higher level of variability for women compared to men.

Index Terms— Forensics, face recognition, data analysis

1. INTRODUCTION

Automatic face recognition over forensic caseworks is still a challenge for the research community. Large amounts of research are being carried out trying to compensate variability sources (such as illumination, pose, facial expressions, occlusions, etc.) that affect significantly reducing the performance of the face recognition systems. In a forensic scenario, these variability factors are crucial, due to frequently forensic examiners have to deal with face images extracted from CCTV cameras and other low quality sources, which makes the task really difficult.

Many different techniques have been developed to automatically tag facial landmarks on a face [1, 2, 3, 4]. These techniques achieve good results over good quality and frontal faces, but are still not good enough for the cases of having high variability and low quality images. Actually, in practice there is still no automatic system that can achieve such a detail level compared to humans. On the other hand, humans are subjective and do not work as systematically as computers. For this reason, in practice forensic examiners make use of semiautomatic systems, which can help in the suspects identification tasks [5].

Among the tasks carried out by forensic examiners, they analyse the intra-variability of two face images, a set of gallery images (with known identity) and the probe image. In an anthropometric analysis they extract manually a set of facial landmarks, then compute some distances between them, which can be used as features in their analysis. Figure 1 shows a diagram of this procedure.

This paper focuses on an analysis of the variability of facial landmarks in a forensic scenario over a database of mugshots. This landmarking variability is affected by two factors, on the one hand the accuracy of the process of landmark tagging, which can vary significantly due to the quality of the images, and on the other hand it is also affected by the intrinsic variation of the landmarks, due to changes in pose, expression or occlusions among others. In this paper, we are interested in studying this variability having images taken at three different distances (3, 2, and 1 meters) between the camera and the persons. A database of mugshot images from 50 persons has been acquired following indications from the Spanish Guardia Civil [6]. 21 facial landmarks were defined and 1200 images have been manually tagged imitating the work performed by a forensic examiner. Some of the findings of this study are that in general facial landmarks located in the outer part of the face (highest point on the head, ears and chin) have a high level of variability, due possibly to hair occlusions. Regarding the distances between the camera and the persons, the variability increases gradually with the distance but not very much. The findings of this paper could be included in the work carried out by a forensic examiner within a anthropometrical facial analysis.

The remainder of the paper is organized as follows. Section 2 describes the acquisition of the mugshot database comprised of 50 persons and 3 distances between the camera and the persons. Section 3 describes the task of landmark tagging and the image processing. Section 4 describes the experimental results achieved and finally Section 5 draws the final conclusions and future work.

2. DATABASE COLLECTION

This section describes the mugshots database collected to carry out the experimental work reported in this paper, using the pro-

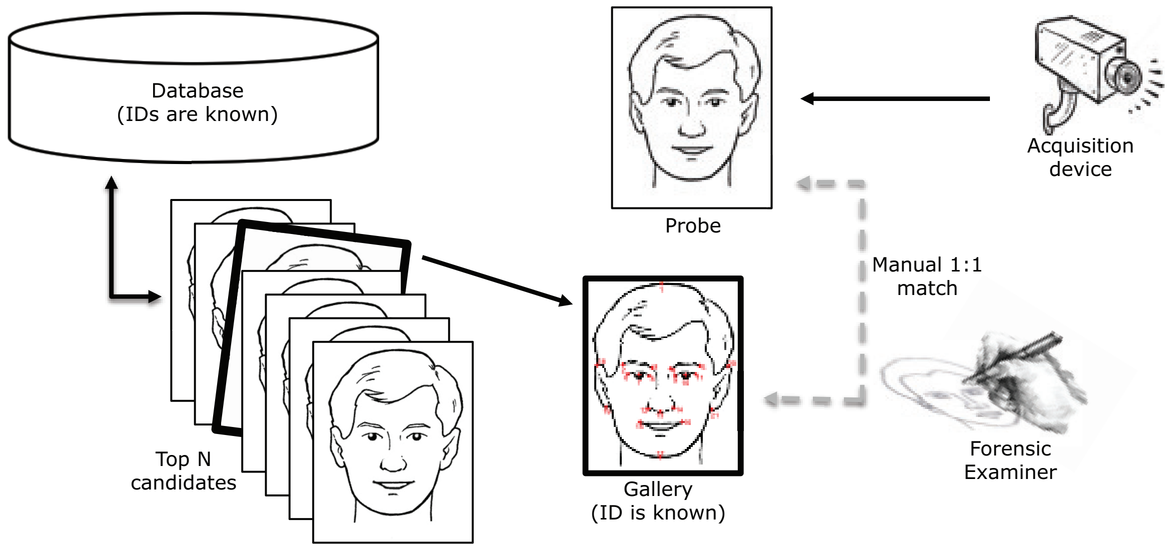


Figure 1: General procedure followed by a forensic examiner to compare two face images.

tocols followed in practice by the Spanish Guardia Civil. Six pictures of each person were taken at every mugshot photo session:

1. Full body. 2 pictures: front and lateral right view. Approximately three meters distancie from camera to person.
2. Upper body. 1 picture: front view. Approximately two meters distancie from camera to person.
3. Face. 3 pictures: front, lateral right (+90 degrees) and semi-lateral left (-45 degrees). Approximately one meter distancie from camera to person.

In this work, only the three frontal images (full body, upper body and face) have been used to analyse the variability of facial landmarks. An example of the three frontal images captured is shown Figure 2. The camera used in the database collection was a Canon EOS 400D and, as a can be seen in the figure, a vertical meter was used in the background, as done in practice.

The database comprises data from 50 persons (32 men and 18 women) acquired in two different sessions. The collection process took place from July to November 2012. The sessions were collected in different days for the same persons. In each session the procedure was repeated four times. Therefore, obtaining a total of 1200 face images (50 persons $\times$ 2 sessions $\times$ 4 times $\times$ 3 face images).

(a) Far (b) Medium (c) Close

Figure 2: Examples of front images acquired at a mugshot session considering three distances between the person and the camera.

3. FACIAL LANDMARK MANUAL TAGGING AND IMAGE PROCESSING

This section describes the process of manual facial landmark tagging and image processing in order to analyse the variability of facial landmarks.

The first step after database collection was to define a set of facial landmarks to include in this study. A set of 21 facial landmarks was defined following recommendations from the Spanish Guardia Civil [6], Netherland Forensic Institute [7] and ENFSI [8], including the irises (2 landmarks) , inner and outer eye corners (4), eyebrow ends (4), mouth corners (2), nose corners (2), center of the nose (1), chin (1), upper and lower ears ends (4) and highest point on the head (1). Figure 3 shows the 21 facial landmarks considered in this paper.

A manual tagging of the 21 landmarks for the whole database was carried out by the same person, imitating the work of a forensic examiner. This manual tagging was performed over the acquired mugshot images.

Then, a second stage of image processing was carried out in order to normalise the facial images to the same size and position. Thus, the midpoint between the eye corners (midpoint between points 6 and 8, and midpoint between 9 and 11) was computed and used instead of the irises positions to align the faces, because the position of the irises can vary if the person does not look at the camera directly. The positions of these two points were fixed having 75 horizontal pixel between them following the recommendation from the ISO standard [9]. Therefore, translation, rotation and scaling of the original images was carried out to normalize the database. This was done in the same way for images collected at different distances between the camera and person. Figure 4 shows an example of the three face images shown in Figure 2 but size normalised, and showing the positions of the 21 facial landmarks in red and the positions of the center of the eyes in green.

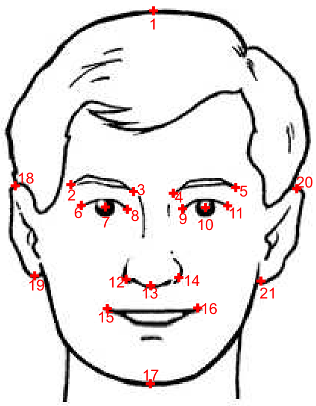


Figure 3: Defined 21 facial landmarks manually tagged.

(a) Far (b) Medium (c) Close

Figure 4: Same example as in Fig. 2 but normalizing the faces with 75 pixels between the center of the eyes. Also, the 21 facial landmarks are shown (red), plus the center of the eyes (green).

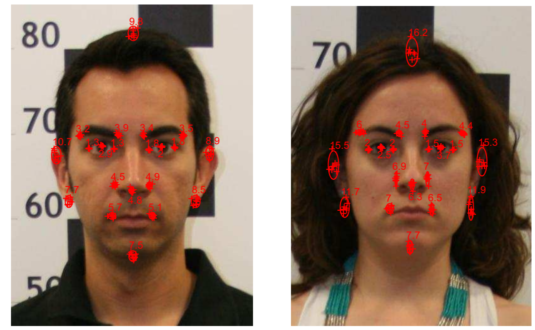
4. EXPERIMENTAL RESULTS

This section describes the experimental work carried out to analyse the variability of facial landmarks considering a forensic controlled scenario. Three different experiments were designed and are described here.

4.1. Person Specific Landmarking Variability

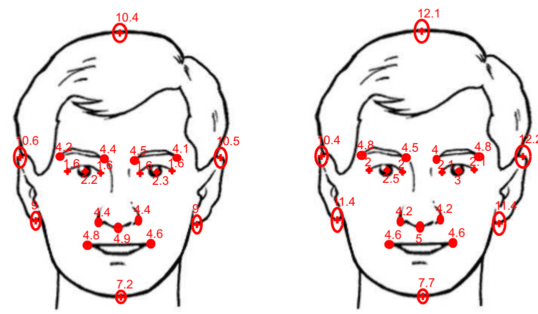
In this experiment a person specific landmarking variability (LV) was studied. Thus, the 8 available facial images per person and per distance were considered. The mean and standard deviation for each facial landmark were computed for the two (x,y) spatial dimensions ($\sigma_{x,i}$, $\sigma_{y,i}$, with *i=1,...,21*), assuming following a gaussian distribution. Figure 5 shows an example face image superimposing for each facial landmark the result of tagging the 8 available images. An elipse around each facial landmark is computed using as the radios $2\sigma_{x,i}$, $2\sigma_{y,i}$. Throughout this paper the variability of the different facial landmarks was computed as $\pm$ 2$\times$mean($\sigma_{x,i}$,$\sigma_{y,i}$), covering this way a 95.44% of the distribution. For example, in the image shown in Figure 5(a) the landmark for the highest point on the head shows a variability of $\pm$ 9.8 pixels considering a normalization of the face images with 75 horizontal pixels between the eye positions.

This procedure was followed for the 50 persons comprising the database, and it was found that the variability of the facial landmarks, specially for the outer ones varies significantly from person to person. Figure 5 shows an example of two persons for the case of the 3 meter distance between camera and person. The variability of these landmarks on the outer part of the face (highest point on the head, chin and ears) is very dependent on



(a) (b)

Figure 5: Examples of the landmarking variability for two persons present in the database for mugshots taken at 3 meters distance between the person and the camera.



(a) Male (b) Female

Figure 6: Results of the landmarking variability for male and female for pictures taken at 3 meters distance between the subjects and the camera.

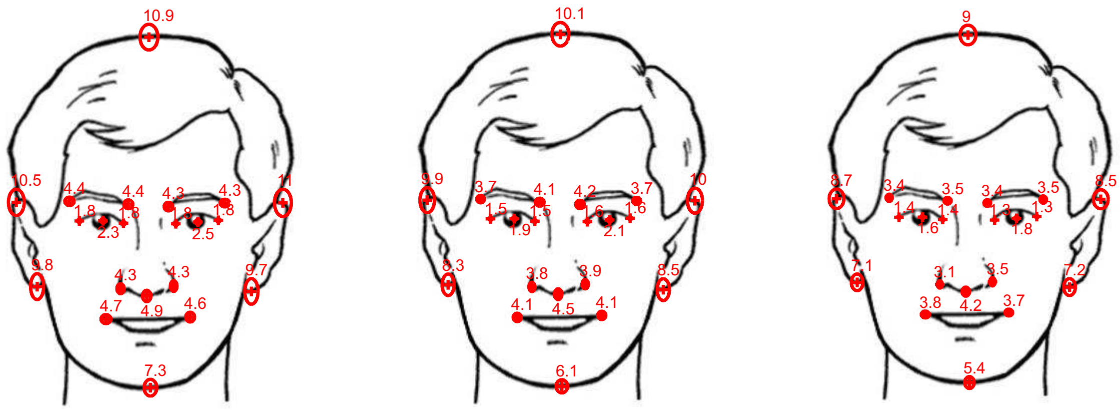
hair occlusions, more frequent in women than men for the population considered. Therefore, the second experiment is designed to study the variability of facial landmarks across gender.

4.2. Gender Specific Landmarking Variability

This section describes the experimental results of the variability of the facial landmarking comparing results achieved for males and females contained in the database. In order to compute a global landmarking variability (LV) for males and females, the mean of the different individual values of the variability of each facial landmark is computed, following the equation:

$$LV_i = \frac{1}{N}\sum_{j=1}^{N}(\sigma_{x,i,j} + \sigma_{y,i,j}) \qquad (1)$$

where *i=1,...,21* are the landmarks and *j=1,...,N*, being N the maxima number of males or females respectively. Figure 6 shows the results achieved for male and female respectively for the case of the 3 meter distance mugshot. As can be seen, in general the landmarking variability is greater for female compared to male (16 landmarks out of the 21), mostly for the landmarks placed in the outer part of the face (i.e., highest point on the head and ears), where there can be more possible hair occlusions. Still the difference in absolute number of pixels having normalised the images with 75 pixels between the eyes is not very significant.



(a) Far (b) Medium (c) Close

Figure 7: Results of the landmarking variability for the three distances considered between the persons and the camera: far (3 meters), medium (2 meters) and close (1 meter).

4.3. Distance Specific Landmarking Variability

This section reports the experimental results achieved for the global landmarking variability (i.e., without distinction for males and females) considering the effect of the 3 distances between the camera and the persons. For each of the three distances, the global landmaking variability is computed in the same way as per Equation 1, but in this case with N being the total number of of persons contained in the database (50).

Figure 7 shows the results achieved. As can be seen, in general the landmarking variability increases with the distance between camera and person (the case with highest variability is the far case). But statistically the variability of the different facial landmarks is very consistent, so we can conclude that in the close scenario, as the images are originally much bigger (around 500 pixel between the eyes), the manual tagging can be done with more precision, and therefore reducing a bit the landmarking variability.

Another finding is that, as already mentioned, the outer parts of the face present more variability (highest point on the head, ears and chin), then mouth and nose areas, and the parts with the least variability are the eyes and eyebrows. It is worth noting that the normalisation of the faces was done using the center of the eyes, so it is also natural that these parts present less variability than the rest.

5. CONCLUSIONS

This paper reports an study of the variability of facial landmarks over a mugshot database collected following the procedures of the Spanish Guardia Civil. Mugshots are taken with three distances between the persons and the camera (3, 2, 1 meters) showing the full body, the upper body and the face respectively. 21 facial landmarks were defined and the database was manually tagged imitating the procedure followed by a forensic examiner. The main conclusions are that the landmarks located in the outer part of the face have a much higher variability compared to the landmarks placed near the eyes. A reason for this is mainly that these are the areas which can have hair occlusions more frequently like the highest point on the head and the ears. This is more accentuated for the females.

Regarding the distances between the camera and the person, the landmarking variability increases with the distance, but not very significantly. In the case considered here, images taken at 3 meters have around 100 pixels between the center of the eyes, which is not a very small size. In case of having mugshots taken for further distances we believe that the landmarking variability would be much higher.

For future work we propose to compare these results with an automatic landmark tagging approach in order to check if these findings are consistent in both approaches. Also, we propose to carry out a face recognition experiment using the landmark positions for both manual and automatic systems and compare results.

6. ACKNOWLEDGEMENTS

R. Vera is supported by a Juan de la Cierva Fellowship from the Spanish MINECO. This work has been partially supported by a contract with Spanish Guardia Civil and projects BBfor2 (FP7-ITN-238803), Bio-Challenge (TEC2009-11186), Contexts (S2009/TIC-1485), TeraSense (CSD2008-00068) and "Cátedra UAM-Telefónica".

7. REFERENCES

[1] S. Gupta, M. K. Markey, and A. C. Bovik, "Anthropometric 3D Face Recognition". Int. Journal of Computer Vision, Vol.90, Num.3, pp.331-349, December 2010.

[2] S. Arca, P. Campadelli, and R. Lanzarotti. "An Efficient Method to Detect Facial Fiducial Points for Face Recognition". In Proc. of ICPR'04, Vol. 1. pp. 532-535, 2004.

[3] S. Arca, P. Campadelli, and R. Lanzarotti. "A Face Recognition System based on Automatically determined Facial Fiducial Points". Pattern Recognition, Vol.39, Num.3, pp.432-443, March 2006.

[4] G. M. Beumer, Q. Tao, A. M. Bazen, and R. N. J. Veldhuis."A Landmark Paper in Face Recognition". In Proc. of Int. Conference on Automatic Face and Gesture Recognition (FG'06), pp. 73-78, USA, 2006.

[5] A. K. Jain, B. Klare and U. Park, "Face Matching and Retrieval in Forensics Applications". IEEE Multimedia, Vol.19, Num.1, pp.20, January 2012.

[6] "Spanish Guardia Civil", http://www.guardiacivil.es/.

[7] "Netherlands Forensic Institute (NFI)", http://www.forensicinstitute.nl.

[8] "European Network of Forensic Science Institutes, http://enfsi.eu/.

[9] ISO/IEC JTC 1/SC 37 N 504, "Biometric Data Interchange Formats Part 5: Face Image", 2004.