Forensic Science International 233.1-3 (2013): 75 – 83

DOI: http://dx.doi.org/10.1016/j.forsciint.2013.08.020

Identification using Face Regions: Application and Assessment in Forensic Scenarios

Pedro Tome, Julian Fierrez, Ruben Vera-Rodriguez and Daniel Ramos

Biometric Recognition Group - ATVS, Escuela Politecnica Superior
Universidad Autonoma de Madrid
Avda. Francisco Tomas y Valiente, 11 - Campus de Cantoblanco - 28049 Madrid, Spain
{pedro.tome, julian.fierrez, ruben.vera, daniel.ramos}@uam.es

Abstract

This paper reports an exhaustive analysis of the discriminative power of the different regions of the human face on various forensic scenarios. In practice, when forensic examiners compare two face images, they focus their attention not only on the overall similarity of the two faces. They carry out an exhaustive morphological comparison region by region (e.g., nose, mouth, eyebrows, etc.). In this scenario it is very important to know based on scientific methods to what extent each facial region can help in identifying a person. This knowledge obtained using quantitative and statical methods on given populations can then be used by the examiner to support or tune his observations. In order to generate such scientific knowledge useful for the expert, several methodologies are compared, such as manual and automatic facial landmarks extraction, different facial regions extractors, and various distances between the subject and the acquisition camera. Also, three scenarios of interest for forensics are considered comparing mugshot and Closed-Circuit TeleVision (CCTV) face images using MORPH and SCface databases. One of the findings is that depending of the acquisition distances, the discriminative power of the facial regions change, having in some cases better performance than the full face.

Keywords:
Forensics, facial regions, facial components, primary biometrics, face recognition.

1. Introduction

Automatic face recognition has been extensively researched over the past two decades. This growth is due to its easy acquisition and its important role in a growing number of application domains, including access control, video surveillance, and its wide use in government issued identity documents (e.g., passport and driving license) [1].

An area where these kinds of systems have obtained an increased emphasis is the forensic field [2]. Forensic science analyses data collected by law enforcement agencies in order to prove or disapprove the guiltiness of a suspect with high confidence under the legal system.

While DNA and fingerprint forensic identification are two of the most reliable and available identification methods in forensic science, automatic face recognition technology still needs to improve the set of available tools to determine a person's identity, particularly from video surveillance imagery. Such progress for forensic face recognition is one of the goals of the FBI's Next Generation Identification program [3].

Face recognition in video surveillance scenarios is a very challenging task due to the variability that can be present. In this sense, there are several studies [4, 5, 6, 7] based on realistic scenarios trying to understand the effect of the different variability factors in this field.

Automatic face recognition systems are generally designed to match images of full faces. However, in practice, forensic examiners carry out a manual inspection of the face images, focussing their attention not only on the full face but also on individual traits. They carry out an exhaustive morphological comparison, analysing the face region by region (e.g., nose, mouth, eyebrows, etc.), even examining traits such as marks, moles, wrinkles, etc.

There are some previous works where region-based face recognition is studied [8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18] but non of them focus their attention in

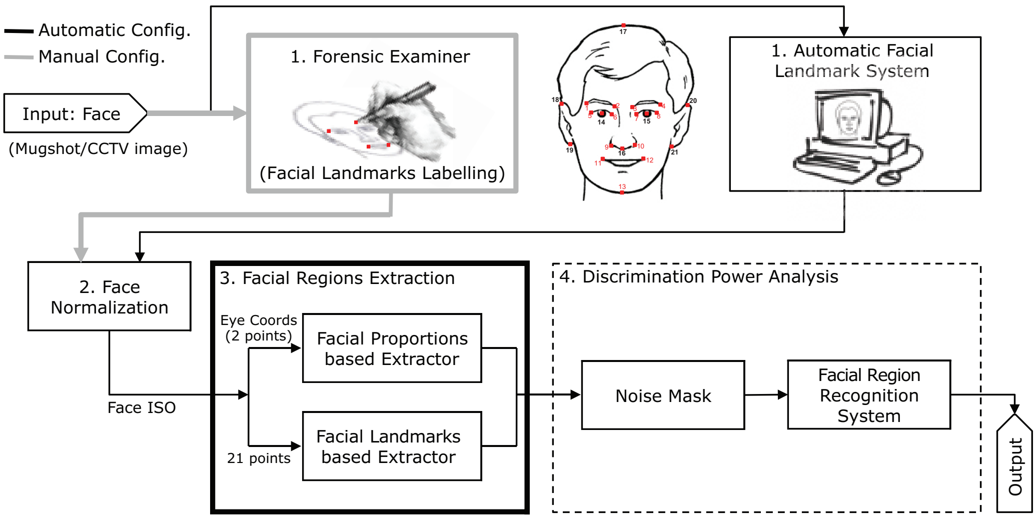

Figure 1: Experimental framework followed to study the discrimination power of the 15 facial regions.

the regions normally considered by forensic experts. In this work, we have extracted facial components (called from now on facial regions) following forensic protocols from law enforcement agencies, allowing us to study the discriminative power of different facial regions individually. In particular, we address in this paper the problem of finding the most discriminative areas of the face for recognition on different acquisition scenarios.

Understanding the discrimination power of different facial regions on a wide population has some remarkable benefits, for example: *i*) allowing investigators to work only with particular regions of the face, *ii*) preventing that incomplete, noisy, and missing regions degrade the recognition accuracy. Further, a better understanding of the individuality of facial regions should facilitate the study of facial regions-based face recognition. In the same way that the field of cognitive science continues to investigate the precise roles of facial regions and holistic processing in human face perception [19], automatic face recognition algorithms also need to explore the role that facial regions processing could have improving their performance.

The main contribution of this paper is: −exhaustive quantitative analysis of the discriminative power of different forensic facial regions extracted from a human face following the procedures of forensic examiners at different distances between the subject and the camera.

Additionally, we propose and compare a standard framework for facial regions extraction useful for controlled and uncontrolled scenarios based on face proportions and facial landmarks with manual or automatic tagging. Fig. 1 shows a diagram of the methodology followed in this paper.

The remainder of this paper is organized as follows. Section 2 provides an overview of the automatic facial region extraction procedure. Section 3 presents the databases used and the feature extraction and classification adopted for the experiments. Section 4 explains the experimental protocol followed and Section 5 presents the experimental results achieved. Finally, Section 6 draws some conclusions of our work.

2. Facial Regions Extraction

This section describes the experimental framework developed to extract the 15 different facial regions considered in this work, following three steps:

1. Detection of facial landmarks.
2. Face normalisation and facial region extraction.
3. Representation of facial regions using eigen-regions.

2.1. Facial Landmark Detection

The first step is to extract a predefined set of anthropometric landmarks. This step has two different configurations: automatic and manual in order to find the facial landmarks.

Given the variability of facial appearances, as well as the variability caused by pose and expression changes, the extraction of facial landmarks is often a difficult task to be performed automatically. When considering challenging scenarios at a distance, the low quality of the images introduces an other important factor that makes even more difficult the detection task.

On the other hand, the common practice of forensic examiners is mainly based on manual and individual skills using some general image processing tools. This approach permits to have reliable landmark information even in lower quality images but may introduce a subjective bias in the process.

In this paper we have followed both an automatic and a manual approach for facial landmark detection.

For the automatic approach, the commercial SDK Luxand FaceSDK 4.0 [20], was first used to automatically detect 65 facial landmarks. Next, the landmarks of each facial region (eyebrows, eyes, nose, mouth and chin) were selected and the rest were removed. The result of this step is an initial placement of facial landmarks where just 13 of them are considered as Fig. 2 top shows. These 13 facial landmarks have been selected following forensic face recognition protocols By Spanish Guardia Civil [21] and NFI [22] and they indicate the terminations of each trait in a human face.

For the manual approach for landmark detection, a human manually tagged 21 facial landmarks imitating the procedure of a forensic examiner, as shown in Fig. 2 bottom. As can be seen the 13 automatic facial landmarks are included as a subset of the 21 marked with the manual approach. In the manual approach, the ears and the upper end of the head are also marked.

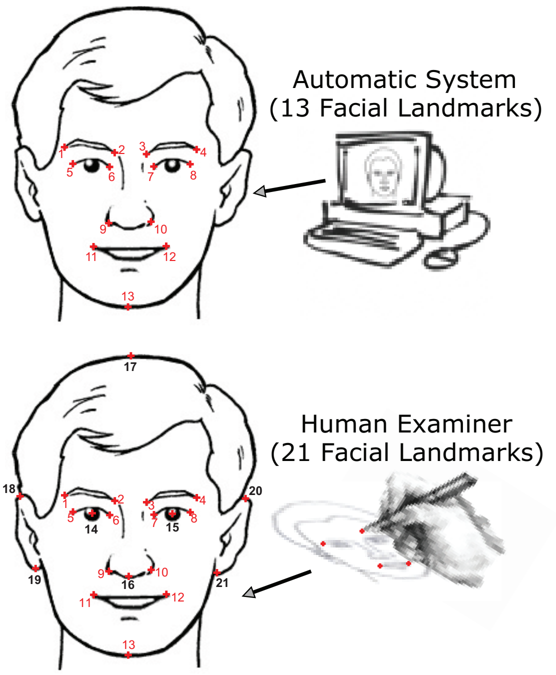

Figure 2: Facial landmarks selected for the automatic and manual configurations.

2.2. *ISO Normalization and Facial Region Extraction*

Once the facial landmarks have been detected, the next step is the extraction of the facial regions. This is performed following two approaches: *i*) based on human face proportions, and *ii*) based on facial landmarks.

Before extracting the facial regions all the faces were normalised following the ISO standard [23] with an interpupillary pixel distance (IPD) of 75 pixels. This step eliminates variations in translation, scale and rotation in horizontal plane, and provides a normalized face in order to compare it or extract facial regions with a standard size for all faces considered.

2.2.1. *Extractor based on Facial Proportions*

The extractor based on facial proportions uses the proportionality relationships in a human face. These relationships divide the human face in several horizontal and vertical areas with the same size as shown in Fig. 3. There are previous works where facial proportions of a human face were studied [24, 25, 26]. Based on these

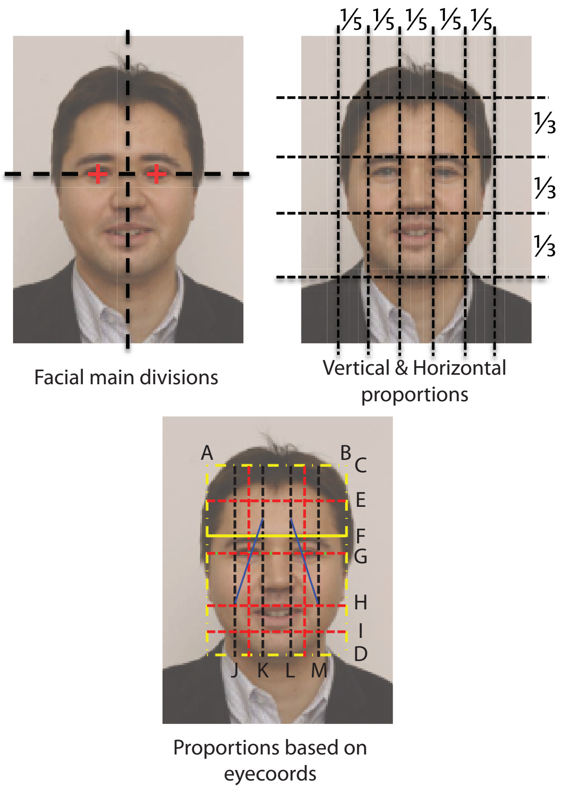

Figure 3: Facial proportions: main facial divisions, horizontal, vertical and proportions based on eyecoords.

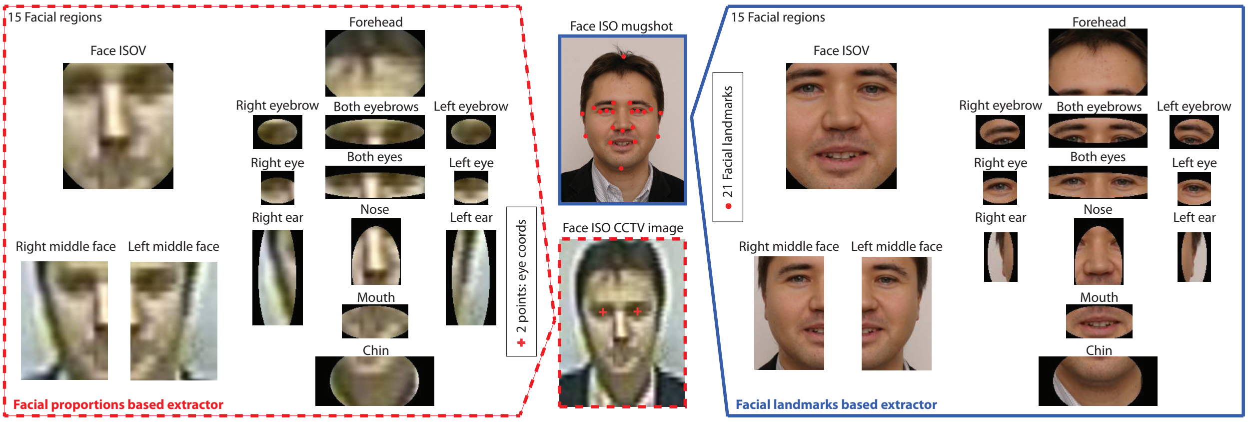

Figure 4: Facial regions extraction. On the left side, with dashed line, the extractor based on facial proportions and on the right side, with solid line, the extractor based on facial landmarks.

Id Num.	Facial Region	Prop. based Extractor	Landmarks based Extractor
1	Chin	55x188	75x181
2	Left ear	145x76	75x51
3	Right ear	145x76	75x51
4	Left eyebrow	48x57	51x75
5	Right eyebrow	48x57	51x75
6	Both eyebrows	48x132	51x151
7	Left eye	48x57	51x51
8	Right eye	48x57	51x51
9	Both eyes	48x132	51x151
10	Face ISOV	192x168	192x168
11	Forehead	71x132	101x151
12	Left middle face	180x132	173x106
13	Right middle face	180x132	173x106
14	Mouth	57x113	51x101
15	Nose	112x76	101x75

Table 1: Facial regions sizes for both extractors based on proportions and facial landmarks (height × width in pixels).

works an automatic facial region extractor system following these proportions rules has been developed [15]. This extractor applies facial proportions rules using the eye centers as reference point. Fig. 3 (right) shows the proportions calculated based on these reference points (lines A-M) used to extract the 15 facial regions described in Fig. 4 (left).

Using just the two eyes coordinates, following single facial proportions rules considering the IPD distance, 15 facial regions (eyebrows, eyes, nose, mouth, etc.) can be approximately extracted from a frontal face. The main drawback of this approach is the low precision, which can produce small misalignments of the region for the different face images.

On the other hand, this extractor would be of interest in challenging uncontrolled scenarios where landmarks are very difficult to be extracted automatically, but an automatic face recognition system can locate the eyes coordinates easily or they can be tagged manually. An example of this extraction can be seen in Fig. 4 (left), which shows the 15 regions considered based on protocols from international forensic laboratories [21, 22].

2.2.2. *Extractor Based on Facial Landmarks*

The second extractor, is based on anthropometric facial landmarks allowing us to extract the facial regions with higher precision.

In this case, a facial region is extracted by estimating the center between each one of two facial landmarks per facial trait and by applying a vertical and horizontal offset to generate a bounding box that contains the facial region, as can be seen in Fig. 5. This procedure is followed automatically for the extraction of the 15 forensic facial regions as shown in Fig. 4 (right).

The main drawback of this approach is that the precision of the extraction depends on the correct manual or automatic localization of the facial landmarks. On the other hand, this method provides a good alignment allowing us to compare facial regions keeping their relationships of shape and size.

There are previous techniques [27, 18] which have used pre-defined cropping boundaries, and a more recent work [9] uses alignment approaches such as Procrustes analysis [28]. In our case, the ISO normalization step previously applied, together with the central point estimation step allows us to solve alignment problems in the extraction process.

Table 1 shows the size of the 15 facial regions for the two extractors. As can be seen, the extractor based on proportions needs a bigger bounding box than the extractor based on facial landmarks.

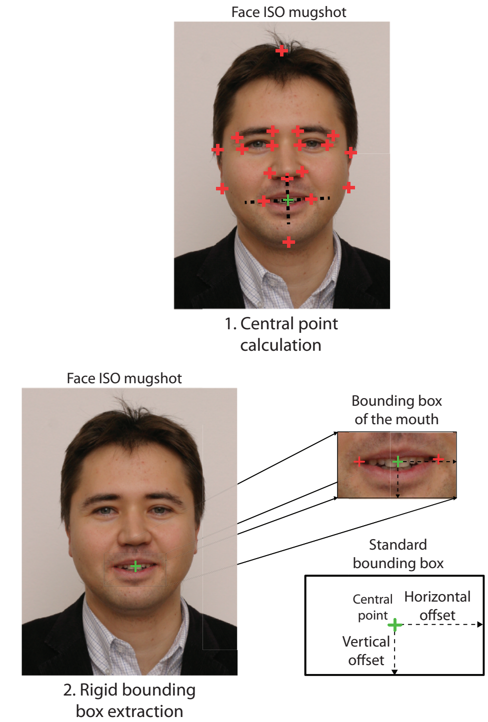

Figure 5: Extraction procedure of the mouth region using the extractor based on facial landmarks.

As it will be seen in the experiments, depending on the scenario at hand, one approach can be superior to the other.

The experimental framework implemented based on these two extractors allows the extraction of 15 different facial regions as can be seen in Fig. 4. The election of these 15 regions is based on protocols from international forensic laboratories [21, 22].

2.3. Representation of Facial Regions

Once each facial region has been aligned and extracted, eigen-regions (Principal Component Analysis, PCA) from each facial region are computed as this type of feature is very popular in face recognition [29].

Each face image is first divided into the regions described in Table 1. Then histogram equalization is applied to each grayscale facial region. In order to avoid external noise in each region, a noise mask is applied (see black areas in Fig 4). Then, eigen-region (PCA) is applied to each facial region over the training set considering the first 200 principal components for each region. Similarity scores are computed in this PCA vector space using a Support Vector Machine (SVM) classifier with a linear kernel [15].

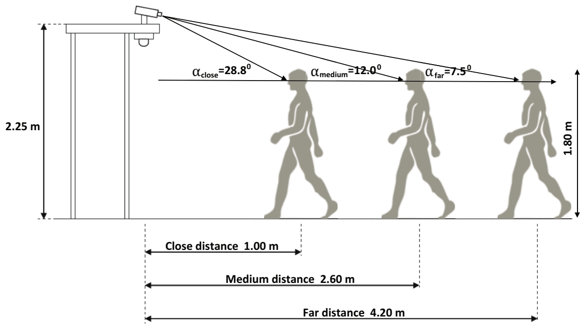

Figure 6: SCface database. There are three different acquisitions distances: *close*, *medium* and *far*. Acquisition angle of each distance calculated for a subject with mean height of 1.80 meters.

3. Databases

The experimental work described in this paper has been carried out using a collection of mugshot and CCTV face images of 130 subjects from two different databases: SCface [30] and MOPRH [31].

SCface is a database of static images of human faces with 4.160 images (in visible and infra-red spectrum) of 130 subjects.

The database is divided into 6 different subsets: *i*) mugshot images, which are high resolution frontal images, and *ii*) five visible video surveillance cameras. Each of these subsets contains 130 images, one per subject. As shown in Fig. 6 the images were acquired in an uncontrolled indoor environment with the persons walking towards several video surveillance cameras having different qualities. Additionally the images were acquired at three different distances: 4.20m (*Far*), 2.60m (*Medium*) and 1.00 meters (*Close*) respectively while the subject walked towards the cameras. Fig. 7 (top) shows an example of a mugshot image, and the images acquired by one of the surveillance cameras. As can be seen there is a considerable scenario variation in terms of quality, pose and illumination. The effect of the pose is specially important due to the different angles between the person and the cameras as shown in Fig. 6.

This database is of particular interest from a forensic point of view because images were acquired using commercially available surveillance equipment, under realistic conditions.

One of its drawbacks is that it is just comprised of one mugshot session, so the comparison of mugshots versus mugshots cannot be carried out only being able to compare mugshots vs. CCTV images and CCTV vs. CCTV.

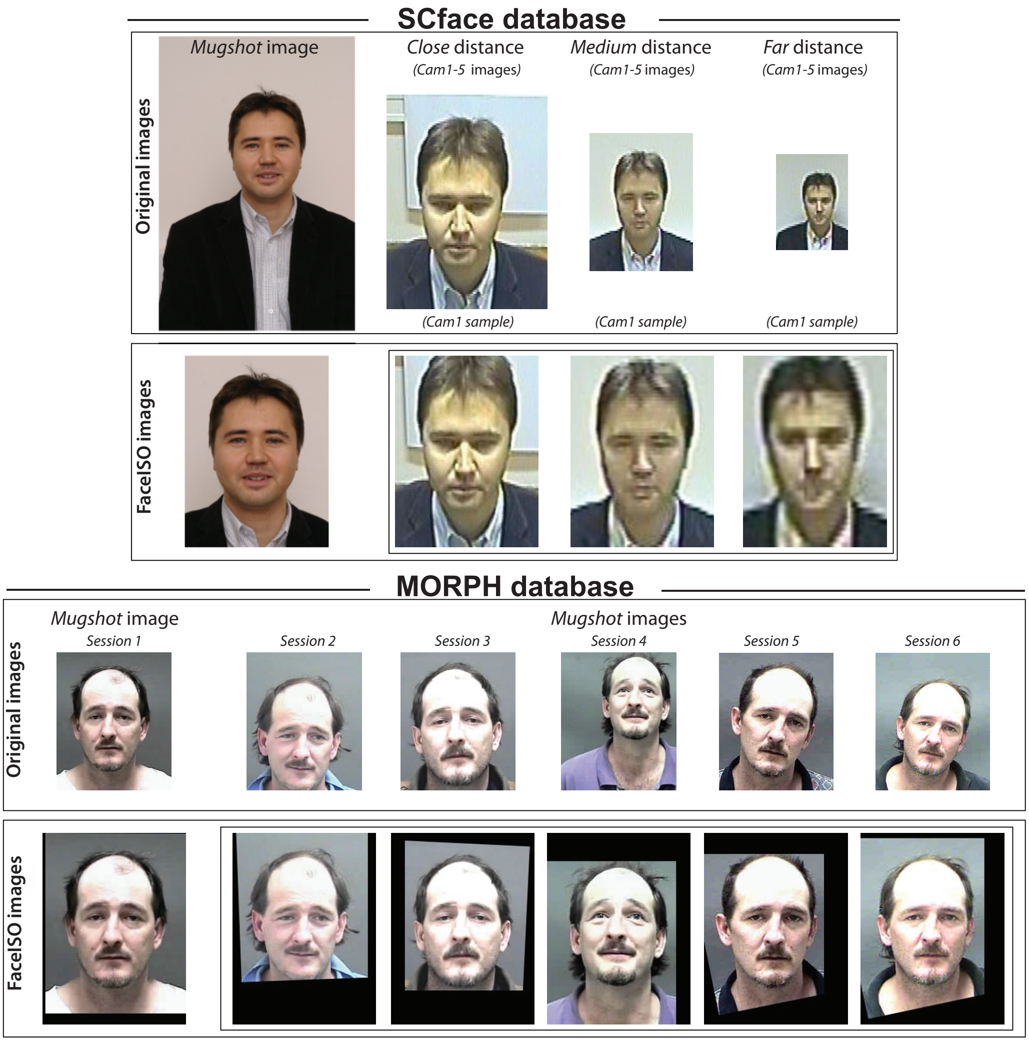

Figure 7: (Top) SCface image samples of each datasets for *mugshot* and *Cam*1 images, and their corresponding normalized face ISO for the *close*, *medium*, and *far* distance. (Bottom) MORPH image samples (200×240) of each session and their corresponding normalized face ISO (300×400).

As a second dataset for our experiments we used the MORPH Non-Commercial Release database [31]. MORPH corpus contains 55.000 frontal face images from more than 13.000 subjects, acquired from 2003 to late 2007. The distribution of ages ranges from 16 to 77 with an average age of 33. The average number of images per individual is 4 and the average time between pictures is 164 days, with the minimum being 1 day and the maximum being 1.681 days. The MORPH database is divided in 5 subsets: *i*) African, *ii*) European, *iii*) Asian, *iv*) Hispanic and *v*) Other.

The subset "European" comprised of 2.704 subjects (2.070 males plus 634 females) was selected for our experiments. Fig. 7 (bottom) shows an example of the images available for a person of the database. We generated a new dataset comprised of 780 mugshot images for 130 subjects with 6 sessions per subject and with a time lapse between sessions around one year. As a result, a similar structure compared to SCface DB is obtained, which facilities their comparison.

It is important to note that SCface was collected in 5 days while the time lapse in MORPH database is around one year. This is an important difference that will be considered in the experimental results and findings.

4. Experimental Protocols

The experimental protocol followed in this paper is similar to the one proposed in [32][1]. The database was divided into 3 subsets based on the subject ID: development (1-43), SVM training (44-87), and test (88-130).

In this work three different protocols are defined considering the different cases that a forensic examiner can find in practice:

1. *Mugshot vs mugshot* protocol
2. *Mugshot vs CCTV* protocol
3. *CCTV vs CCTV* protocol

These three protocols are considered to extract conclusions that can be helpful in the forensic practice or for improving the traditional face recognition systems in these challenging scenarios.

In addition, three distances between subject and camera are analysed: *close*, *medium* and *far* distance. The analysis of these 3 configuration is also of great interest for forensics and face biometrics.

4.1. Mugshot vs Mugshot

This protocol has been defined in order to study the performance of different facial regions in a controlled scenario. For this particular case the subset of the MORPH database previously described is used.

This protocol compares good quality mugshot images against the same kind of images. The development set consists of the 6 available images per subject (1 image × 6 sessions) for 43 subjects (218 images in total, subjects 1 to 43). This set is used to train the PCA subspace.

Each subject model in the Test set (subjects 88 to 130) is then trained using the first session (s1), as Client data for SVM Training and all images from subjects 44 to 87 as Impostor data. As test images we consider the other 5 sessions (s2-s6) in the Test set. This dataset partitioning is summarized in Table 2.

[1] http://scface.org/

	MORPH DB (130 Subjects) - Mugshot vs Mugshot protocol		
Subsets	1...43 Subject (43 Subjects)	44...87 Subject (44 Subjects)	88...130 Subject (43 Subjects)
*s*1	Development set (PCA subspace)	SVM Training (Impostors)	SVM Training (Clients)
*s*2			Test (Clients/Impostors)
*s*3			
*s*4			
*s*5			
*s*6			

Table 2: Partitioning of the MORPH DB according to the *Mugshot vs Mugshot images* evaluation protocol.

	SCface DB (130 Subjects) - Mugshot vs CCTV protocol		
Subsets	1...43 Subject (43 Subjects)	44...87 Subject (44 Subjects)	88...130 Subject (43 Subjects)
Mugshot	Development set (PCA subspace)	SVM Training (Impostors)	SVM Training (Clients)
Cam 1			Test (Clients/Impostors)
Cam 2			
Cam 3			
Cam 4			
Cam 5			

Table 3: Partitioning of the SCface DB according to the *Mugshot vs CCTV images* evaluation protocol.

4.2. Mugshot vs CCTV

This scenario is common in forensic laboratories, and it is very challenging because the difficulty in finding reliable similarities between doubted CCTV images and undoubted mugshot images from police records. For this reason, the results obtained in this scenario are specially helpful for the forensic practice.

In this case each subject model is trained using a single mugshot image (SVM Training Clients). Then, test images are taken from the 5 surveillance cameras at 3 different distances: *close*, *medium* and *far* (Test set). The Development and SVM Training sets are similar to the previous protocols as can be seen in Table 3.

4.3. CCTV vs CCTV

A third protocol was designed to compare CCTV against CCTV images. In this case the same variability factors (low resolution, pose, illumination, etc.) affect both train and test images (see *Cam1* images in Fig. 7 (top)). This protocol was defined in order to understand the performance when the training set is influenced by the same variability factors present in test images.

As shown in Table 4, the partitioning of the SCface DB into Development, SVM Training, and Testing is similar to the previous protocols, only considering in this case the information from the 5 surveillance cameras, and using the first one for modelling each subject (through SVM Training).

	SCface DB (130 Subjects) - CCTV vs CCTV protocol		
Subsets	1...43 Subject (43 Subjects)	44...87 Subject (44 Subjects)	88...130 Subject (43 Subjects)
Cam 1	Development set (PCA subspace)	SVM Training (Impostors)	SVM Training (Clients)
Cam 2			Test (Clients/Impostors)
Cam 3			
Cam 4			
Cam 5			

Table 4: Partitioning of the SCface DB according to the *CCTV vs CCTV images* evaluation protocol.

In this scenario the system is trained with images with *close* distance and compared with images from the three distances: *close*, *medium*, and *far*.

5. Experimental Results

This section describes the experimental results and findings achieved following the protocols described in Sect. 4. The main goal of the experiments is to study the discrimination power of the different facial regions.

First, a comparison between manual and automatic facial landmark detection is described in Sect. 5.1. Then, results for the three experimental protocols described in Sect. 4 are presented.

5.1. Comparison of Manual and Automatic Facial Landmark Detection

This experiment analyses the error introduced in the process of automatic facial landmark tagging with respect to manual tagging (ground truth).

Fig. 8 shows the normalised average error in number of pixels for the 13 facial landmarks considered. Results are computed for the different datasets considered in this paper.

As can be seen there is a notable difference between the mugshot subsets for the MORPH and SC-face databases (i.e., error is much higher in MORPH mugshot). This difference is due to the original image size (resolution), MORPH images are 240×200 pixels, and SCface mugshot images are 3072×2048. Facial landmark tagging over high resolution images can be much more accurate compared to lower resolution images [33].

CCTV images have a higher error compared to SC-face mugshots. We have to note that the image size for the *close* (224×168), *medium* (144×108), and *far* (100×75) scenarios is different, which may be the main reason for the increasing error between *close* and *far*.

It is interesting to note that the landmarks for CCTV on the right part of the face image present a higher error compared the left side. This effect can be due to illumination or pose artifacts. The mugshot images do not present the previous effect.

As a result, we observe an increasing error between mugshot and CCTV, and between *close* and *far* distances for the automatic landmark detection compared to the labelling done by a human expert used as ground truth.

5.2. Mugshot vs Mugshot

This section presents the results for the mugshot versus mugshot scenario using the MORPH database. Results for both manual and automatic landmark tagging together with the two facial region extractors are presented and compared.

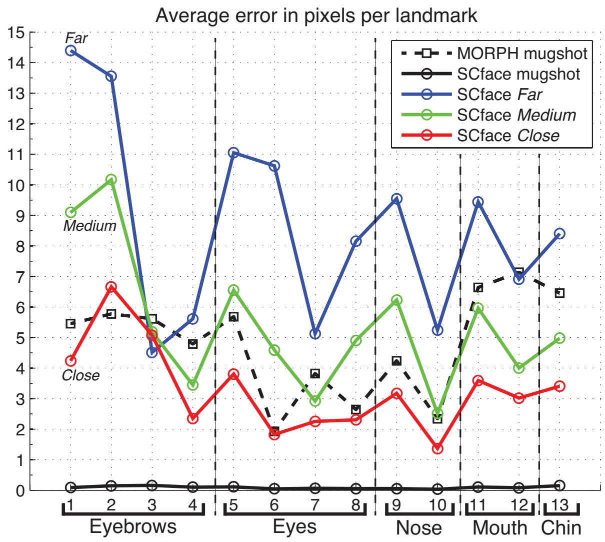

Figure 8: Comparative error analysis between the automatic facial landmark system and a manual examiner (ground truth) based on Euclidean distance in the different scenarios analysed. Pixel values are normalised to 240×200 image size. Legend of landmark's number can be seen in Fig. 2.

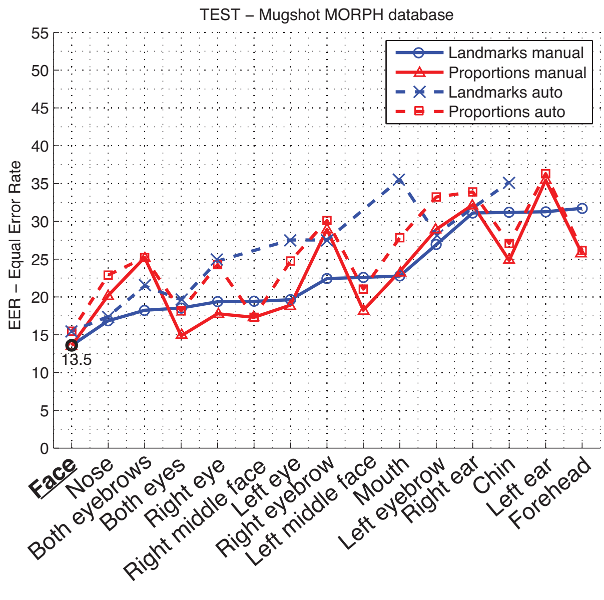

Figure 9: EER values for the different facial regions extracted for the mugshot vs mugshot images scenario. Curves are ordered by the manual landmarks results.

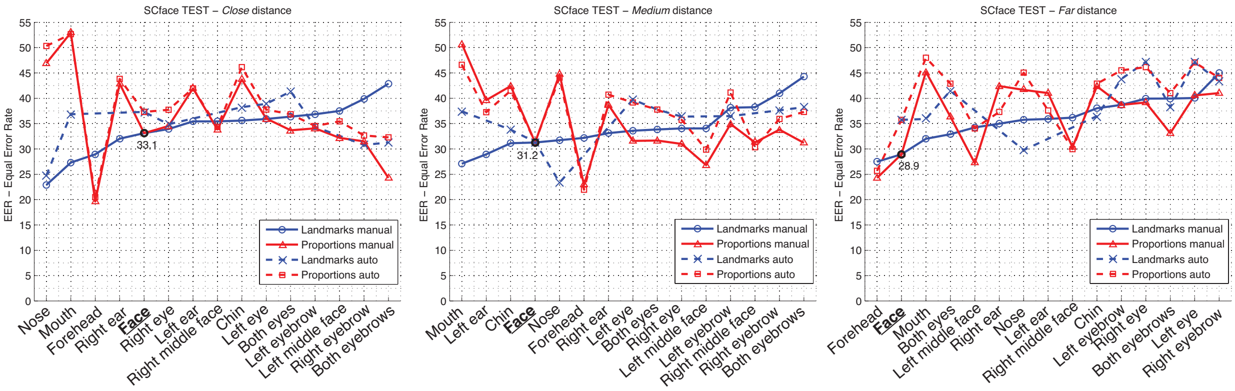

Figure 10: EER values for the different facial regions extracted for the three different distances: *close*, *medium* and *far* for the mugshot vs CCTV images scenario.

Better results are obtained with the manual landmark tagging in both extractors as shown in Fig. 9. This graph presents the Equal Error Rate (EER) [1] of the whole *face* region compared with the rest of the facial regions extracted. The 15 facial regions are ordered from lower to higher EER (left to right).

The *face* region achieves the best recognition performance. Inner traits of the face such as the *nose*, *both eyebrows*, *both eyes*, etc., have better performance in this mugshot controlled scenario compared to the outer traits of the face such as the *ears*, *chin*, and *forehead*. This is in concordance with previous works [15, 33, 34, 16].

Regarding the two facial region extractors, it is interesting to see that some regions achieve a better performance for the extractor based on the proportions (*both eyes, middle faces, chin*, and *forehead*), and some others achieve a better performance using the extractor based on the landmarks (*nose, both eyebrows, mouth,* and *ears*).

5.3. Mugshot vs CCTV

As discussed before, this is probably the most interesting and challenging scenario for forensic examiners.

Results are shown in Fig. 10, where EER achieved for each facial region extracted is represented over the three scenarios at a distance: *close*, *medium* and *far* for the SCface database.

Fig. 10 presents a very interesting experimental finding comparing the EER of the *face* region with the rest of the facial regions extracted. As can be seen, the recognition performance considering the whole face improves when the distance increases (31.1% to 28.9% EER for *close* and *far* distance, respectively). We believe this is mainly due to the varying acquisition angle. As can be seen in Fig. 6 this angle is smaller in the *far* scenario, and therefore the pose is more similar to the mugshot image. This can also explain that the best facial region performance is achieved for the *nose, mouth* and *forehead* for the three distances (*close*, *medium*, and *far*) respectively.

Similar to the previous scenario, here better results are obtained in all scenarios at a distance by the manual landmarks tagging for both extractors, as could be expected.

Another very interesting finding is the discriminative power achieved by the *forehead* region in the *far* distance which is better than the full *face*. These results could be due to the system used where the features based on PCA may be not representing well the full potential of mugshot and surveillance camera images.

From a global point of view, both extractors based on proportions and facial landmarks experiment different performances for the different regions, but it is interesting to see how they follow the same trend for the three distances considered (i.e., manual are usually better than automatic landmarks, and proportions-based is usually better than landmark-based region extraction). This effect could be due to that the extractor based on proportions estimates the facial regions approximately, which even using the rigid noise masks, can include more information and improves the EER.

These results suggest that for low quality images at a distance, facial regions could be extracted just using two points (eye coordinates), which an automatic face recognition system can locate easily and the recognition result of each region would be similar to the one obtained using a more sophisticated facial landmark extractor.

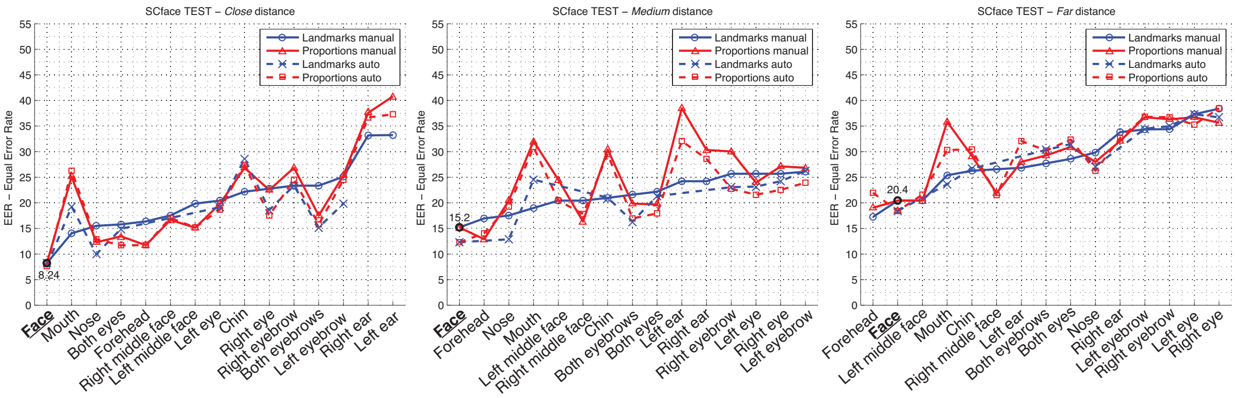

Figure 11: EER values for the different facial regions extracted for the three different distances: *close*, *medium* and *far* for the CCTV vs CCTV scenario.

5.4. *CCTV vs CCTV*

This scenario presents better performances in all distances than the analysed before on the SCface database. This improvement could be because the training and testing sets include more or less the same environmental variability (see *Cam1* images in Fig. 7 (top)), and the effect of the acquisition angle (pose) is not so important considering that all cameras are in the same static position and never totally frontal as the acquisition of mugshot images. Here, the system is trained with images from the *close* scenario and compared with images from the three scenarios: *close*, *medium*, and *far*.

It is important to emphasize the unexpected by good performance achieved in this scenario, which is better thanthe one in the *mugshot vs mugshot* scenario. This result can be explained by the much larger time lapse between training and testing data for MORPH DB (mugshot vs mugshot) compared to SCface (CCTV vs CCTV).

Results are shown in Fig. 11 where a similar tendency between both extractors under study is observed. In particular it is very remarkable the similar performance for the *far* scenario. This demonstrates that, the proposed simple region extractor based on only eye coordinates and face proportions could be very useful for unconstrained scenarios at a distance.

In general terms there is a decrement of performance when the distance increases. In this case, as expected because training and testing conditions are similar here, the performance of the *face* region decreases with the distance between the subject and the camera.

The best performances are achieved for the *face, forehead, nose* and *mouth* in the three distances. Again the *forehead* region reaches the best performance in the *far* scenario and the second position in the *medium* distance scenario. This could be because 115 subjects of the database are male and just 15 are female. Male subjects usually have short hair and therefore the forehead is free of occlusions. Female subjects on the other hand, usually have long hair and more occlusions which may lead to decreased performance in this region. The *forehead region* reaches an important role in this uncontrolled scenario in comparison with the previous *mugshot versus mugshot* scenario. While in controlled scenarios the *forehead* region achieved the worst results, here this facial region outperforms the other facial regions. Even in the *medium* and *far* scenarios this facial trait is one of the most discriminative.

6. Conclusions

This paper reports an exhaustive analysis of the discriminative power of the different regions of the human face on various forensic scenarios.

We first described an experimental framework to extract 15 different facial regions of a human face following forensic protocols with 4 variants: manual or automatic landmark detection, and then region extraction based either on the full set of landmarks (13 and 21 for automatic and manual landmark detection, respectively) or only the eye coordinates and general face proportions.

The comparison of the two region extractors resulted in better recognition performance for the outer facial regions when using the extractor based on proportions. In contrast better recognition performance is achieved for the inner facial regions when using the extractor based on landmarks. As a result, we obtain that the extractor based on proportions can be very useful in scenarios at a distance, where obtaining reliable landmark information with automatic systems is very difficult. Also interestingly, similar performance is obtained with both extractors in the *far* scenario where images are degraded. This means that for low quality images at a distance, facial regions could be extracted just using two points (eye coordinates), and the recognition result of each region would be similar to a facial landmark extractor.

Differences between manual and automatic annotations of facial landmarks have been also analysed. When the acquisition distance increases, the error for the automatic approach considering the manual labels as ground truth also increases.

After analysing the landmark and region extraction phases,we studied three scenarios with different distance between subject and camera common in forensic casework. In all cases, we obtain that the recognition performance of facial regions depends on the acquisition distance. The best three facial regions with high discrimination power in the *close* distance are the *face, nose*, and *forehead*. However in *far* distance, the best performance is achieved by the *forehead*. This facial region acquires an important role on scenarios at a distance such as CCTV versus CCTV. It was noted that this result could be due to the great majority of short hair males, as with females that region may be much more reliable.

In the most common forensic casework (*mugshot versus CCTV* images), variability factors have a high importance and produce a decrement of recognition performance with respect to the more controlled *mugshot vs mugshot* scenario.

In addition as being useful background information that can guide and help experts to interpret and evaluate face evidences, these findings can have a significant impact on the design of face recognition algorithms.

References

[1] A. K. Jain, A. A. Ross, K. Nandakumar, Introduction to Biometrics, Springer Science+Business Media, LLC, 2011.

[2] A. K. Jain, B. Klare, U. Park, Face recognition: Some challenges in forensics., in: FG, IEEE, 2011, pp. 726–733.

[3] Next Generation Identification, http://www.fbi.gov/about-us/cjis/fingerprints_biometrics/ngi/ngi2.

[4] X. Zhang, Y. Gao, Face recognition across pose: A review., Pattern Recognition (2009) 2876–2896.

[5] P. Tome, J. Fierrez, M. C. Fairhurst, J. Ortega-Garcia, Acquisition scenario analysis for face recognition at a distance, in: 6th International Symposium on Visual Computing (ISVC), 2010.

[6] S. Z. Li, A. K. Jain (Eds.), Handbook of Face Recognition, 2nd Edition. Springer, 2011.

[7] P. Tome, R. Vera-Rodriguez, J. Fierrez, J. Ortega-Garcia, Variability compensation using NAP for unconstrained face recognition, in: Proc. 10th International Conference on Practical Applications of Agents and Multi-Agent Systems (PAAMS12), Vol. 151, Springer, 2012, pp. 129–139.

[8] A. K. Jain, B. Klare, U. Park, Face matching and retrieval in forensics applications, IEEE MultiMedia 19 (1) (2012) 20.

[9] K. Bonnen, B. Klare, A. K. Jain, Component-based representation in automated face recognition, IEEE Transactions on Information Forensics and Security 8 (1) (2013) 239–253.

[10] M. Tistarelli, Active/space-variant object recognition, Image and Vision Computing 13 (3) (1995) 215–226.

[11] B. Heisele, T. Serre, T. Poggio, A component-based framework for face detection and identification, Int. J. Computer Vision 74 (2) (2007) 167–181.

[12] F. Li, H. Wechsler, M. Tistarelli, Robust fusion using boosting and transduction for component-based face recognition, in: 10th International Conference on Control, Automation, Robotics and Vision, ICARCV 2008, Hanoi, Vietnam, 17-20 December 2008, Proceedings, IEEE, 2008, pp. 434–439.

[13] S. Gupta, M. K. Markey, A. C. Bovik, Anthropometric 3d face recognition, Int. J. Comput. Vision 90 (3) (2010) 331–349.

[14] T. Ali, P. Tome, J. Fierrez, R. Vera-Rodriguez, L. Spreeuwers, R. Veldhuis, A study of identification performance of facial regions from cctv images, in: 5th International Workshop on Computational Forensics, IWCF 2012, 2012.

[15] P. Tome, L. Blazquez, R. Vera-Rodriguez, J. Fierrez, J. Ortega-Garcia, N. Exposito, P. Leston, Understanding the discrimination power of facial regions in forensic casework, in: International Workshop on Biometrics and Forensics, Lisboa, Portugal, 2013.

[16] O. Ocegueda, S. K. Shah, I. A. Kakadiaris, Which parts of the face give out your identity?, in: Proccedings of CVPR, IEEE, 2011, pp. 641–648.

[17] J. Sadr, I. Jarudi, P. Sinha, The role of eyebrows in face recognition, Perception 32 (3) (2003) 285–293.

[18] T.-K. Kim, H. Kim, W. Hwang, J. Kittler, Component-based lda face description for image retrieval and mpeg-7 standardisation, Image Vision Comput. 23 (7) (2005) 631–642.

[19] J. Gold, P. Mundy, B. Tjan, The perception of a face is no more than the sum of its parts, Psychol Science 23 (4) (2012) 427–34.

[20] Luxand Face SDK, http://www.luxand.com.

[21] Spanish Guardia Civil (DGGC), http://www.guardiacivil.es/.

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

[23] ISO/IEC 19794-5:2011, Information technology – biometric data interchange formats – part 5: Face image data, in: International Organization for Standardization, 2011.

[24] T. Shiang, A statistical approach to data analysis and 3-d geometric description of the human head and face., Proc Natl Sci Counc Repub China B 23 (1) (1999) 19–26.

[25] Y. Jefferson, Facial beauty-establishing a universal standard, IJO 15 (1) (2004) 9.

[26] H. Gunes, M. Piccardi, Assessing facial beauty through proportion analysis by image processing and supervised learning., International Journal of Man-Machine Studies 64 (12) (2006) 1184–1199.

[27] K. Pan, S. Liao, Z. Zhang, S. Li, P. Zhang, Part-based face recognition using near infrared images, in: Computer Vision and Pattern Recognition, 2007. CVPR '07. IEEE Conference on, 2007, pp. 1–6.

[28] J. Gower, Generalized procrustes analysis, Psychometrika 40 (1) (1975) 33–51.

[29] M. Turk, A. Pentland, Face recognition using eigenfaces, in: Computer Vision and Pattern Recognition, 1991. Proceedings CVPR '91., IEEE Computer Society Conference on, 1991, pp. 586–591.

[30] M. Grgic, K. Delac, S. Grgic, Scface - surveillance cameras face

database, Multimedia Tools Appl. 51 (3) (2011) 863–879.

[31] K. Ricanek, T. Tesafaye, Morph: a longitudinal image database of normal adult age-progression, in: Automatic Face and Gesture Recognition, 2006. FGR 2006. 7th International Conference on, 2006, pp. 341 –345.

[32] R. Wallace, M. McLaren, C. McCool, S. Marcel, Inter-session variability modelling and joint factor analysis for face authentication, in: Proceedings of the 2011 International Joint Conference on Biometrics, IJCB '11, IEEE Computer Society, Washington, DC, USA, 2011, pp. 1–8.

[33] R. Vera-Rodriguez, P. Tome, J. Fierrez, N. Exposito, F. J. Vega, Analysis of the variability of facial landmarks in a forensic scenario, in: International Workshop on Biometrics and Forensics, Lisboa, Portugal, 2013.

[34] T. Faltemier, K. Bowyer, P. Flynn, A region ensemble for 3-d face recognition, Information Forensics and Security, IEEE Transactions on 3 (1) (2008) 62–73.