BIOFACE, a Biometric Face demonstrator

Mourad Ouaret
Orange, France-Telecom R&D
Sophia Antipolis, France
mourad.ouaret@orange-ftgroup.com

Antitza Dantcheva, Rui Min
Eurecom
Sophia Antipolis, France
{antitza.dantcheva,rui.min}@eurecom.fr

Lionel Daniel, Jean-Luc Dugelay
Eurecom
Sophia Antipolis, France
{lionel.daniel,jean-luc.dugelay}@eurecom.fr

SUBMITTED to ACM MULTIMEDIA 2010 DEMO PROGRAM

ABSTRACT

In this paper, a demonstrator called BIOFACE incorporating several facial biometric techniques is described. It includes the well established Eigenfaces and the recently published Tomofaces techniques, which perform face recognition based on facial appearance and dynamics, respectively. Both techniques are based on the space dimensionality reduction and the enrollment requires the projection of several positive face samples to the reduced space. Alternatively, BIOFACE also performs face recognition based on the matching of Scale Invariant Feature Transform (SIFT) features.

Moreover, BIOFACE extracts a facial soft biometric profile, which consists of a bag of facial soft biometric traits such as skin, hair, and eye color, the presence of glasses, beard and moustache. The fast and efficient detection of the facial soft biometrics is performed as a pre-processing step, and employed for pruning the search for the facial recognition module.

Finally, the demonstrator also detects facial events such as blinking, yawning and looking-away. The car driver scenario is a good example to exhibit the importance of such traits to detect fatigue.

The BIOFACE demonstrator is an attempt to show the potential and the performance of such facial processing techniques in a real-life scenario. The demonstrator is built using the C/C++ programming language, which is suitable for implementing image and video processing techniques due to its fast execution. On top of that, the Open Source Computer Vision Library (OpenCV), which is optimized for Intel processors, is used to implement the image processing algorithms.

General Terms

Algorithms, Security and Verification.

Keywords

Face Recognition, Soft Biometrics and Facial Event Detection.

1. INTRODUCTION

Biometrics consists of methods to uniquely recognize humans based on physical or/and behavioral traits. Out of these traits, face is one of the most characteristic and discriminative human body parts for a person. Consequently, researchers study image processing techniques to extract robust facial traits for person identification. The interest for such mostly security-related technologies has increased lately, due to the apparent growth of criminality rate and terrorist threats as well as the availability and the cost efficiency of video sensors. A variety of applications motivate and utilize identification/authentication techniques in biometrics. There is also the will of governments to improve the security conditions of their citizens that encouraged the deployment of such systems.

Intuitively, the first attempts were to recognize persons based on facial information. Eigenfaces [1] is one of the essential basic techniques for person recognition by using facial appearance. It is based on the notion of dimensionality reduction by applying the Principal Component Analysis (PCA) [2]. Most of face recognition techniques rely on the appearance (the actual pixel values). On the other hand, Tomofaces is introduced in [3] as a novel technique for face recognition based on the discrete video tomography [4] to characterize the head's dynamics. It applies a temporal X-ray transformation to a video to summarize the facial motion. Both techniques require the computation of a reduced space out of the enrollment data, which is mostly performed offline. On the other hand, face recognition based on SIFT [5] features performs both, recognition and enrollment, online in real time. In this case, SIFT extracts feature points from both, reference and query, images and matches them by comparing the local descriptors. The same person appearing in both images should yield several matching points. The matching process's complexity increases with the number of extracted points and the size of the computed descriptors.

Soft biometrics entails the advantages of fast, computational efficient and enrollment-free biometric analysis, even in the absence of cooperation and consent of the person. These advantages can improve a combined facial recognition system in terms of robustness, computational time and resources. Former work on soft biometrics includes a recent method for human head detection based on hair-color in [6], a robust eye glasses detection method [7], a beard removal algorithm [8], and finally an extensive survey of skin color detection methods [9]. Jain examined soft biometric traits in a recent work [10], where detection of facial marks accompanies classical face recognition. Treatment of eye color detection as a facial soft biometric trait constitutes, to our knowledge, a new research path. Note however, that iris pattern and texture have been widely studied, not taking eye color information in consideration.

According to the recent report of the American department of Transportation [11], driver fatigue is recognized as one of the critical vehicle safety issues and has received the attention from both laboratory researchers and car industry. A review of the driver fatigue monitoring research [12] shows that the majority of state-of-art methods exploit the information captured from a single video camera, where one or multiple facial events are detected in order to assist in making decision for fatigue detection. The Viola-Jones object detection framework [13] has been considered as a benchmark in nowadays computer vision community. Objects like face, eye and mouth can be detected under this framework using trained classifiers. The training is performed on a large number of positive and negative samples. Taking advantage of this existing technology, we have made a preliminary attempt at facial events detection for driver monitoring systems.

The paper is organized as follows. Initially, the foundations of the implemented image processing algorithms are reviewed from sections 2 to 4. Then, the demonstrator's hardware requirements and its Graphical User Interface (GUI) are described in section 5. Finally, some concluding remarks are drawn in section 6.

2. FACE RECOGNITION

In this section, the basics of Eigenfaces and Tomofaces are reviewed as they constitute important modules of our biometric application. Tomofaces is based on a temporal X-ray transformation of a video sequence. This transformation consists in a canny edge detector [14] applied to each frame of the video sequence, containing the head and the shoulders of the person. Afterwards, the temporal video X-ray transformation consists in adding up the edge frames over time and a scaling, which adjusts the X-ray image upper range values. The resulting X-ray image is filtered to attenuate its brightest background contours that could result from a static textured background. This is performed by cutting down to zero the X-ray image values superior to the background attenuation threshold. An example of the resulting X-ray image is depicted in Figure 1.

Figure 1 Example of an X-ray image.

Then, a PCA is applied to the X-ray images to isolate the discriminative information that characterizes the individual in terms of dynamics. Further, the projections of the X-ray images into the reduced space for each person are averaged (centroid) to compute its model. Finally, the identification is achieved for a queried X-ray image by nearest neighbor classification.

For eigenfaces, it is exactly the same procedure as tomofaces except that the input in this case is no longer an X-ray image but grayscale versions of the video frames.

SIFT is an alternative to the previously mentioned techniques to perform face recognition. In this case, SIFT extracts characteristic points from the query face image. These points are invariant to scale and orientation, and are highly distinctive of the face. Then, they are matched to the ones extracted from the face image of the person when enrolled to the system. The matching is based on the descriptors comparison and the coherence of the matched points' locations. In other words, matched points should belong to the same region of the face since the matched objects are human faces with the same orientation. This is achieved by looking for a match within a window around the original position. Then, a score is computed for the query image, which depends on the number of matches and their locations. More specifically, the face is divided into three regions, the inner, outer and in between regions. Matched points in the inner region of the face image have a weight of 3. On the other hand, they have a weight of 1, if they are located in the outer face region. In this case, they most probably belong to the background or the head's contour. Finally, matched points located in between the inner and outer regions have a weight of 2. Finally, the final score is computed as the sum of weights of all the matches.

3. SOFTBIOMETRICS

The demonstrator detects the different soft biometrics traits using the location of the eyes, mouth, nose and upper face. For higher accuracy in the training step all coordinates were manually annotated. We consider a temporal sliding window of multiple frames, which increases the reliability of the detection, without increasing the computational time. The results of the detection are accompanied by a confidence factor for each decision.

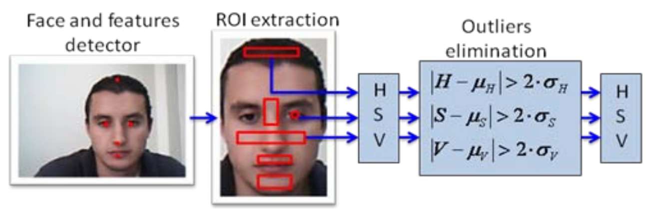

Figure 2 Regions of interest for facial soft biometrics.

The considered regions of interest for the featured facial soft biometric traits are illustrated in Figure 2.

3.1 Glasses detection

We propose for glasses detection an algorithm consisting of the following steps: eye detection, grey-level conversion of the image, histogram equalization, and extraction of the region between the eyes, Laplacian edge detection and finally line detection as the decision carrying part.

3.2 Color-based softbiometrics

We employ for the color based soft biometrics eye, hair and skin color the HSV colorspace, due to the independence between channels and low illumination variance. The HSV values for the respective regions of interest (ROI) were extracted and averaged over the ROIs. In a further step, the outliers were disregarded (all values bigger than $2.\sigma$, σ being the standard deviation of the HSV average value).

We utilized for eye color detection the Hough circle transform to locate the iris, followed by the extraction of the pupil and eventual reflections.

Hair and skin color detection ROI in the upper head is defined and two colors are determined out of possible colors. The latter are modeled by means of Gaussian Mixture Models, whose parameters are estimated through the EM algorithm.

The following soft biometric trait instances can be detected:

Presence of glasses: 0 / 1

Presence of beard: 0 / 1

Presence of moustache: 0 / 1

Eye color: blue / green / gray / brown / black / mixed

Skin color: 1 / 2 / 3

Hair color: blond / brown / black / red / grey/ white / bald

3.3 Moustache and beard detection

We compute three dimensional distances between the color values of the ROIs (above and underneath the mouth, for moustache and beard respectively) and the previously determined skin and hair colors. The minimum Euclidean distance ascertain the decision of existing beard and moustache respectively

4. FACIAL EVENT DETECTION

Furthermore, BIOFACE is able to detect facial events such as blinking, yawning and looking-away as described in the following subsections.

4.1 Blinking detection

The blinking detection implemented in BIOFACE is an improved version of the method presented in [15]. In our facial events detection framework, the blinking detection is triggered as a consequence of the frontal face detection. After the frontal face is detected, the eye detection is performed in the upper part of the face region to locate eye(s). In practice, the eye detection does not always detect both eyes at the same time due to illumination effects. Therefore, we utilize only one eye for blinking detection: The smaller eye is used for blinking detection when both eyes are detected. This does not sacrifice the performance since blink occurs always on both eyes at the same time.

Once the eye is located in the first frame, an eye template (Figure 3) is created to track the eye using template matching. The tracking mechanism is faster and more robust than detecting eye(s) in each frame of the input video. Our system considers blinking as a local facial motion, thus we apply motion detection based on frame difference to detect blinking. First the binary map of image difference between frames of the eye region is computed. Then, a median filter is applied to the binary map to filter out the connected components which are too small and isolated. Finally, if only one connected component appears in the middle of the eye region, one blink is detected.

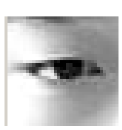

Figure 3 Eye template.

4.2 Yawning detection

The yawn detection in our system is based on the detection of the mouth. The basic idea is when a person is yawning his mouth cannot be detected by the mouth detector trained by the ordinary mouths. Similar to blinking detection, we first use a frontal face detector to approximately locate the mouth position in order to reduce the false detections. Then we apply mouth detection to search a restricted area of the lower part of the face. If the detected mouth obeys some pre-defined heuristic rules (proper size and position), the driver is not yawning. Otherwise, a yawn is detected. In addition, we use a motion estimation scheme to further filter out the yawn candidates based on the fact that a yawn usually causes a head motion.

If the mouth detector is trained properly, both closed mouth and mouth with slight motion (e.g. talking or smiling) can be successfully detected. In this case, the ambiguity between normal facial expressions and an actual yawn can be avoided.

4.3 Looking-away detection

While driving, a person is not supposed to look a way for a long period. Therefore, looking-away detection is also integrated into our system. The algorithm is rather simple. Two independent face detectors are trained, one for frontal face detection and the other for profile face detection. Once a frontal face is detected, it indicates that the person is in the driving position. Then, when the frontal face is missing and a profile face is present, a looking-away event is detected.

5. BIOFACE

BIOFACE has low hardware requirement: the demonstrator runs on a laptop (T8000 2.4GHz, 2Go RAM) having an ordinary webcam (640x480@24fps). BIOFACE was designed with Microsoft Visual C++, SIFT library and OpenCV 1.1 image processing library, which provides functions such as face/eyes/nose/mouth detectors, contour extractors, PCA calculations.

In this technical demonstration, conference guests could either passively enjoy the soft biometrics and abnormal event module, or actively enroll themselves into the authentication/identification system based on Eigenfaces, Tomofaces or SIFT-based face recognition. The Graphical User Interface (GUI) of the demonstrator is depicted in Figure 4.

6. CONCLUSION

The biometric demonstrator, BIOFACE, for facial image processing is introduced in this paper. It incorporates several algorithms to deal with face recognition, soft biometrics and facial event detection. For face recognition, the following algorithms are supported: Eigenfaces, Tomofaces and SIFT-based face recognition. BIOFACE shows that Tomofaces, which represents the dynamics of the face, adds discriminative information to the appearance-based Eigenfaces. Such face recognition techniques require heavy computations to build the corresponding space and to compute the persons' signatures. On the other hand, the authentication process is less complex and easily performed online. For SIFT-based face recognition, the enrollment is very simple as a snap-shot of the face is captured and stored as a reference picture. On the opposite, the authentication can become much more complex with the number of matched points and the size of the extracted descriptors. Nevertheless, it can be easily simplified by matching each point against the points that lie within a neighboring window. Moreover, the descriptors of the reference image can be extracted offline and stored locally. Thus, descriptors computation is only performed for the query image during the online authentication.

Furthermore, BIOFACE builds a biometric profile for a person by gathering information such as skin, eye and hair color and glasses, beard and moustache presence. Facial events are also detected by the demonstrator for driver monitoring purposes. This can be used to detect fatigue in the car driver scenario for example.

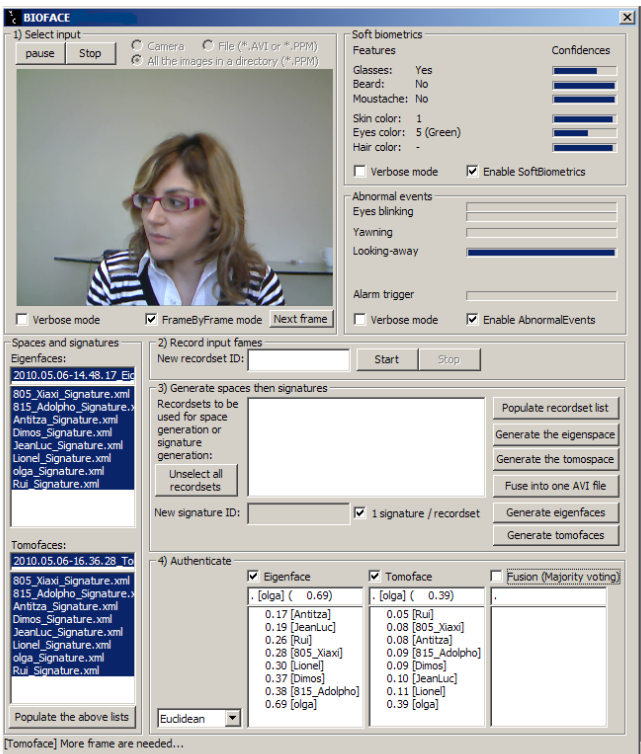

Figure 4 BIOFACE demonstrator GUI.

Unfortunately, there are few technical demos to illustrate and evaluate such techniques under realistic conditions. BIOFACE is a first attempt to exhibit the facial processing algorithms that has been studied and developed within the context of a European project. On the other hand, most published papers report results on databases and the corresponding parameters are tuned to perform well on these specific databases while BIOFACE shows that it possible to tune such techniques to achieve a good performance in a real life scenario. At the same time, it shows the difficulty of some face recognition techniques (Eigenfaces and Tomofaces) adapting to changing conditions (illumination, pose) when compared to others (SIFT).

Future work should aim at evaluating the impact of combining several processing modules on the overall performance. As an example, this can be in the form of a fusion between Eigenfaces and Tomofaces for face recognition instead of having each module working independently.

7. REFERENCES

[1] Turk M.A. and Pentland A.P. ,"*Face recognition using eigenfaces*", IEEE Proceedings on Computer Vision and Pattern Recognition , pages 586-591, June 1991.

[2] Duda R.O., Hart P.E. and Stork D.G., "*Pattern classification*", John Wiley & Sons Interscience, 2nd edition, 2000.

[3] F. Matta J. L. Dugelay, "*Tomofaces: Eigenfaces extended to videos of speakers*", IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Las Vegas, USA, 2008.

[4] Akutsu A. and Tonomura Y, "*Video tomography: an efficient method for camerawork extraction and motion analysis*", ACM Proceedings on Multimedia, pages 349-356, October 1994.

[5] David G. Lowe, "*Object recognition from local scale-invariant features*", International Conference on Computer Vision, Corfu, Greece, pp. 1150-1157, September 1999.

[6] Min Zhao, Dihua Sun, Hengpan He, "*Hair-color Modeling and Head Detection*", World Congress on Intelligent Control and Automation, Chongqing, China, June 25 - 27, 2008.

[7] X. Jiang, M. Binkert, B. Achermann and H. Bunke, "*Towards Detection of Glasses in Facial Images*", *Pattern Analysis & Applications*, Springer London, Vol. 3, pp. 9-18, 2000.

[8] N. Kaushik, and A. Mittal, "*A Novel Beard Removal Method based on Structural Similarity and Co-ordinate Transformations*", Annual IEEE India Conference, pp.1-6, 2006.

[9] P. Kakumanua, S. Makrogiannisa and N. Bourbakis "*A survey of skin-color modeling and detection methods*", Pattern Recognition, Vol. 40, Issue 3, 2007.
WCICA, pp.7773-7776, 2008.

[10] Anil K. Jain and Unsang Park, "*Facial Marks: Soft Biometric For Face Recognition*", International Conference on Image Processing (ICIP), Cairo, Egypt, November 2009.

[11] H. Howarth L. Barr S. Popkin, "*Evaluation of emerging driver fatigue detection measures and technologies. Final report*", John A. Volpe National Transportation Systems Center, Cambridge, MA, June 2009.

[12] A. Williamson, T. Chamberlain, "*Review of on-road driver fatigue monitoring devices*", NSW Injury Risk Management Research Centre, University of New South Wales, [Online], Available: http://www.maa.nsw.gov.au/, April, 2005.

[13] Paul Viola, Michael Jones, "*Robust Real-time Object Detection*", IJCV, 2001.

[14] Canny J., "*A computational approach to edge detection*", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 8, issue 6, pages 679-698, 1986.

[15] http://nashruddin.com/Real_Time_Eye_Tracking_and_Blink_Detection.

[16] http://web.engr.oregonstate.edu/~hess/index.html.

[17] http://opencv.willowgarage.com/wiki/.