https://doi.org/10.2352/EI.2022.34.10.IPAS-355
© 2022, Society for Imaging Science and Technology

Face detection and recognition in organic video: a comparative study for sport celebrities database

Yigit AKBAY, Mihai MITREA
Telecom SudParis, ARTEMIS Department, SAMOVAR Laboratory
19 place Marguerite Perey, 91120 Palaiseau France, mihai.mitrea@telecom-sudparis.eu

Abstract

The present paper reports on an experimental study carried out under the applicative field of organic video processing and related to the possibility of identifying soccer celebrities in video content. In contrast to common state-of-the-art studies, special attention is paid to the cases in which the face is not completely included in the frame (lateral views, partial occlusions, etc.) and/or in which arbitrarily lighting conditions occur. To this aim, after a state-of-the-art study, we consider two conventional types of face detection algorithms (Haar Cascade Classifier, and MMOD – Max-Margin object detection) coupled to two conventional face recognition models (LSBH – Local binary pattern histogram, and CNN-based Pruned ResNet). The experimental work consists of evaluating the performances of the four possible combinations among the above-mentioned two face detection and two face recognition methods. An organic video database of about 1 hour is organized for this study. In addition, a public image database with 31 celebrity's frontal face images is also considered. As an overall conclusion, we brought to light that the MMOD coupled to a Pruned ResNet model seems to better suit the organic video processing use-case constraints, being able to reach an accuracy of 85%. The study also brings to light and discusses the differences in the quantitative results obtained for the two types of databases content (organic video content *vs*. celebrity's face images.

Keywords- Face detection, face recognition, organic video, Haar Cascade Classifier, Max-Margin object detection (MMOD), Local binary pattern histogram (LSBH), Pruned ResNet

1. Introduction

Face detection is the process of determining the areas in an image where a human face is presented, while face recognition is the process of determining a person identity through his/her facial features [1,2]. As humans can spontaneously and seamlessly achieve this task, more and more computer-vision applications try to mimic it: smartphone unlocking, contactless biometric authentication for payments, or fake news detection, to mention but a few. The present paper belongs to the applicative field of organic video that is an emerging trend in Martech (marketing & technologies). It covers the production, distribution and tracking of unpaid and uncontrolled advertising video content. In contrast to generic face recognition applications, the organic video comes across with completely unconstrained content recording and distribution conditions, such as face position relative to the camera, face occlusions (other people nearby, sunglasses), ambient lighting (illumination), video encoding (frame rate, format, codec), …

The paper is structured as follows. After studying the state-of-the-art solutions in Section 2, two face detection and two face recognition methods are selected and considered for the quantitative results in Section 3. Section 3 is structured according to the processed database, to the design of the testbed and to the quantitative results. Conclusions are drawn in Section 4.

2. State of the art

2.1. Face detection methods

The scientific research field of face detection covers a large variety of approaches, from multi-resolution representations to neural networks, passing through information theory or hidden Markov models, to mention but a few. While an exhaustive state-of-the-art study is neither possible nor intended, this section will present the main trends in the field. It is structured according to the taxonomy presented in [3]: knowledge-based methods, feature-invariant methods, template matching, and appearance-based methods. Each of these four classes has the same final applicative purpose, yet their conceptual basis can be structured into an incremental/complementary logic.

First, knowledge-based methods [3, 4] are based on the encoding of the human psycho-cognitive mechanisms as a set of rules expressing the facial features as well as the relationship among and between them. For instance, in [3], the facial features are coded based on human knowledge about the characteristic intensity distribution over facial/non-facial regions. In [4], such relationship is expressed by the skin tone pixels using the chrominance (C_r and C_b) values. Other methods of this type are also reported in [5], [6] and [7]. As it can be noticed, the inner limitation of such approaches is related to the trade-off between the specificity and the generality of these rules (that are directly reflected in the trade-off between false positives and false negatives). From the experimental point of view, such methods are very sensitive to illumination conditions and are expected to be adapted and/or extended to deal with different skin colors.

Secondly, feature-invariant methods [8, 9] are conceptually a binary classifier (facial *vs*. no facial areas in an image). The facial area is detected through morphological elements such as eyes, nose, chin, cheekbones, ears, or forehead. For instance, in [8], the model is first trained as a classifier; thus, it can learn photos of the face and its side angle with pose-invariant representations, effectively recovering photorealistic frontal faces. The Multi-PIE Benchmark, CFP Benchmark, and LFW Benchmark datasets are used in the experiments. The datasets contain different viewpoints and lighting conditions of the face image. Experimental results showed that 99.85% facial recognition accuracy was obtained from the Multi-PIE Benchmark dataset. In [9], a combination of principal component analysis and discrete cosine transform is considered. The experiments are carried out on the Cambridge ORL face database, that consists of 400 images (92x112 pixels, 256 grey levels) representing 40 identities under a high degree of variability in

expression and pose. The experimental results show that the recognition rate can reach 97.5%. The Haar Cascade Classifier faces detection technique is more compatible with real-time applications and allows face detection accuracy of 75.33% to be obtained [10]. Other methods of this type are also reported in [11]. From the methodological point of view, such approaches have as main limitations their variability with the facial expressions.

Thirdly, template matching methods use predefined or parameterized face templates to find or detect faces based on the correlation between predefined or deformable templates and input images. For instance, the study in [12] detects faces in images by calculating the cross-correlation between the specified contrast region (ROI) face in the frame and the model face image. The experiments are carried out on data base created by the authors to this purpose and consider 320 x 240 pixels. The study in [13] considers the skin area selection in the YC_bC_r color space and matching with a reference face template. The method includes the analysis of the DCT energy of the image, complemented by the detection of blinks and pupil movements. The dataset consists of 15 frames (768x1024 pixels). The disadvantage of template matching methods relates to the low detection rate under challenging conditions such as different poses, scales, brightness, and darkness. The advantage is that its application is relatively easy compared to other methods. While such methods are intuitive and low complex (assuming the template is defined), the state-of-the-art experimental results point out to their sensitivity to variations in pose, scale, and shape.

Finally, appearance-based methods can be considered as an incremental step over template-matching methods: they allow the face models (or templates) to be learned from a set of training images that must capture the representative variability of facial appearance, as illustrated in [14], [15], [16]. In [14], an angle based approached is applied on facial features: a face model in the form of a triangular pattern is first created to reflect rules such as the positions and distances of the mouth, eyes, and checks. The dataset consists of 1068 images (250x350 pixels) belonging to 162 persons. A face detection rate of 82.21% was obtained. In [15], face detection is performed using cascading classifiers and Haar-like features on the face image. Pre-processing operations such as dimensional normalization, histogram equalization, and color space transformation are applied to the detected face images. Fisher's Linear Discriminant Analysis, Local Binary Pattern Histograms (LBPH), and Principal Component Analysis methods are used to extract the unique facial features of each face. The experimental results show that the highest performance values (99.39% detection rate) are obtained with the LBPH technique and for 96x96 image. Other methods of this type are also reported in [16], [17], [18]. The advantages of the appearance-based methods are given by their high accuracy as well as by their reduced computation complexity; yet their performances are significantly decreased when processing high-dimensional facial images or on side-viewed facial images. The CNN-based Maximum-Margin Object Detector Model faces detection technique [19] ensures the highest face detection rate and can detect even non-frontal part (half part) of the human face.

This concise state-of-the-art analysis leads us to select for our study the Haar Cascade Classifier [20] technique from the Feature-invariant method category and the CNN-based Maximum-Margin Object Detector [21] model technique in the Appearance-based Method category. These two methods will be considered in conjunction with face recognition methods that will be selected according to the study in Section 2.2.

2.2. Face recognition methods

Spanned over more than the 30 years, face recognition is also a heterogeneous research field that will be illustrated through 5 families of methods: Eigenfaces (1991), Local Binary Patterns Histograms (1996), SIFT – Scale Invariant Feature Transform (1999), SURF – Speed Up Robust Features (2006), and Convolutional Neural Network (2015).

Eigenfaces refer to an appearance-based approach that encodes the face variation in a series of face images and that uses this information to compare images of individual faces in a holistic manner. Specifically, the eigenfaces are the principal components of the distribution of faces, or equivalently, the eigenvectors of the covariance matrix of the set of face images [22], [23], [24].

In [25], the implementation of eigenfaces is achieved through the nearest neighbor (NN) approach to classify the test vectors using the Euclidean distance. Yet, the eigenfaces methods are based on several face descriptors, as for example the PCA – Principal Component Analysis, MPCA – Multilinear PCA or ICA – Independent Component Analysis.

PCA is manly used to model linear variation of high-dimensional data. Its goal is to find a set of mutually orthogonal basis functions that capture the directions of maximum variance in the data and for which the coefficients are pairwise decorrelated [26]. PCA is an unsupervised technique, so the method does not rely on class information. MPCA is an extension of PCA to tensors or multilinear arrays [27], since a face image is most naturally a multilinear array, meaning that there are two dimensions describing the location of each pixel in a face image. In [25], ICA is presented as a generalization of PCA that identifies high-order statistical relationships between pixels to form a better set of basis vectors.

The Local Binary Patterns Histograms (LBPH) do not holistically consider an image but tries to find the local structure of images by comparing each pixel with its neighboring pixels [28]. The comparison process starts by moving a window across the image and at each move (each local part of the picture), the pixel's intensity value is located at the center compared with its surrounding pixels; if the neighbor's value is less than or equal to the center's value, the neighbor value will be set 1, otherwise to 0. After reading these 0/1 values under the 3×3 window in clockwise order, a local binary pattern like 11100011 is generated. When finishing doing this conversion on the whole image, a list of local binary patterns will be generated. After that, converting each binary pattern into a decimal number using binary to decimal conversion and then a histogram of all those decimal values will be generated at the end. In [29] this method was applied to Yale B dataset contain 633 frontal face images. It showed robustness against monotonic grayscale conversions, with average recognition results of 99.74%. Finally, remember that LBPH has computational complexity compatible with real-time applications.

The Scale Invariant Feature Transform (SIFT) is an algorithm designed to detect and describe scale, translation, and rotation-invariant local features in images [30]. The original SIFT algorithm has been successfully applied in a large variety of applications, from general object detection and recognition tasks to panorama stitching. One of its more recent usages also includes face recognition, where it was shown to deliver encouraging results. SIFT-based face recognition techniques found in the literature rely heavily on the so-

called keypoint detector, which locates interest points in the given image that are ultimately used to compute the SIFT descriptors.

The Speeded-Up Robust Features (SURF) is a fast and robust algorithm for local, similarity invariant representation and comparison of images. The main interest of the SURF approach lies in its fast computation of operators using box filters, thus enabling real-time applications such as tracking and object recognition [31].

The use of neural networks for face recognition has been intensively studied, as illustrated in [32-37]: unsupervised, semi-supervised or supervised techniques coexistent. In [32] a model using neural networks and Gabor feature extraction is implemented. The Database of Faces (The ORL Database of Faces) was used in the study and an accuracy rate of 93.33% was obtained. In [33], GA (Genetic Algorithms) and ANN (Artificial Neural Networks) are jointly applied to achieve the fusion and selection of the features required to identify the important areas in the facial region of each person in the frame and to recognize who the face belongs to. In the study, FERET benchmark dataset and its extended version are processed (both datasets represent 50 distinct persons): 94% accuracy has been obtained. In [34], the Eigenfaces technique was used to extract the relevant information to recognize the important features of the human face. Then, using this information, neural network-based face recognition was applied. In the study, 80 face images of 8 people were used and a face recognition rate of 95.4% was obtained. In [35], a face recognition method is introduced by combining the Back Propagation Neural Network (BPNN) technique and Principal component analysis (PCA), in which non-linear face images can be easily recognized. While Principal component analysis (PCA) is used to reduce the dimensionality of the face image, face recognition is done with the Backpropagation Neural Network (BPNN) technique. In the study, 200 face images taken from the Yale dataset were applied and a face recognition accuracy of 87.1% was achieved. Moreover, note that such methods can also be considered in conjunction with earlier approaches: for instance, in [36] a system that uses a combination of eigenfaces and neural network was proposed while [37] stand for a study on coupling PCA and neural network.

This concise state-of-the-art analysis led us to select for our study the LBPH [38] and a CNN [39], [40] method based on a state-of-the-art face recognizer.

3. Experiments

3.1. Experimental database

The experimental results are obtained out of processing two databases.

The first database (the organic video database) is composed of 20 video sequences of about 3 minutes each and is organized by the authors for this study. The content is selected from publicly available content on the Internet and corresponds to the soccer players and coaches, as well as to the first author of the paper, as illustrated in Figure 1 and Figure 2. The video content is encoded at 30 frames/sec and features heterogeneous resolutions, ranging from 360x640 pixels to 1080x1920 pixels. The visual content is heterogeneous, combing excerpts from football games, TV shows, private life events, etc. The capturing conditions are completely unconstrained and are made by both professional and unprofessional cameras.

The second dataset (celebrities face dataset) consists of a total of 2562 frames with 31 different celebrities. The content is selected from the publicly available Face Data of 31 different classes on Kaggle [41]. Images resolutions, ranging from 360x640 pixels to 1340x2010 pixels. In our study, we applied 4 Celebrities which are Jessica Alba, Billie Eilish, Claire Holt, Tom Cruise.

3.2. Experimental setup

The experimental study considers the two conventional types of face detection algorithms (Haar Cascade Classifier, and MMOD) as well as the two conventional face recognition models (LSBH, and CNN-based Pruned ResNet) selected through the state-of-the-art study.

The ML Object Detection Algorithm proposed by Viola and Jones, the Haar Cascade Classifier [10], is considered according to its open-source software implementation [11].

The MMOD Convolutional Neural Networks Model [19] is specifically designed for non-frontal face views captured at variable angles; we considered the open-source implementation available in [17].

LBPH-based face recognition [38] is an ML-based face recognition algorithm considered according to its open-source version available at [42].

To achieve DL-based face recognition, we considered a pruned version of ResNet-34 [40], as implemented at [39]. It is provided by the OpenCV library.

Figure 1. Facial samples images for ML-Based Face recognition dataset (multiple views from the same person are required)

Figure 2. Facial samples images for DL-Based Face recognition dataset

3.3. Performance analysis

To assess the overall performances of the investigate methods, the basic classifier properties are assessed. Both face detection and face recognition can be modelled as binary classifiers: hence, their overall properties can be derived from the confusion matrix that provides the four types of possible results, namely the true positives (*TP*), the false positives (*FP*), the false negatives (*FN*) and the true negatives (*TN*).

TP, *FP*, *FN* and *TN* subsequently allow for the usual classifier properties to be computed, as follows:

- Precision (*Prec*) evaluates the ratio of correct detected images among all positive images:

$Precision = \frac{TP}{TP+FP}$;

- Recall (*Recall*), also referred to as sensitivity, is calculated as the number of correct detected images divided by the total number of positive classified images:

  $Recall = \frac{TP}{TP+FN}$;

- *Accuracy* reflects the correctness of the classifier by the ratio of the total correct answers to the total answers:

  $Accuracy = \frac{TP+TN}{TP+TN+FP+FN}$;

- *Error* is complementary to *Accuracy* and can be computed according to the following formula:

  $Error = \frac{FP+FN}{TP+TN+FP+FN}$.

Note that usually *Prec* and *Recall* are presented by their values (between 0 and 1) while *Accuracy* and *Error* as percentages.

3.4. Experimental results

The experimental results correspond to the successive detection of 12 soccer players and coaches, as well as to the first author of the paper, by using the four possible combinations among the two face detection methods and the two face recognition methods.

The subjective quality evaluation is carried out in collaboration with use-case owners from the Vidmizer (https://vidmizer.com). Figures 3 and 4 illustrate the performances through both simple, frontal frames (Figure 3) and more complex content, combining several faces positioned with arbitrary angles with respect to the camera as well as partial occlusions (Figure 4). Figure 3 shows that for basic tasks (no occlusions, frontal face positions), the four types of methods behave in a quite similar way. Yet, for complex, heterogeneous content, Figure 4 points to different behavior and suggests that MMOD coupled to a Pruned ResNet model would practically outperform the other three combinations.

The objective quality evaluation is carried out in terms of the four entities defined in Section 3.3, namely *Prec*, *Recall*, *Accuracy* and *Error*, and the corresponding results are reported in Figures 5 and 6, for the two databases described in Section 3.1 (organic video, and face detection, respectively). Note that Figures 5 and 6 provide information about the individual face detection and face recognition operations. The end-to-end results are presented in Table 1 for the four investigated configurations and for the two databases.

When using ML based detection techniques on organic video (see Figure 5), the results are quite deceiving: the *Accuracy* and *Error* values are of 45% and 55%, respectively. The use of DL techniques ameliorates these values to 75% and 25%, respectively. When comparing these two sets of results with the ones reported in Figure 6 (i.e. for the second dataset, corresponding to frontal views of celebrities), it can be stated that the low performances are rather related to the type of content than to the methods themselves.

When considering the face recognition operation, the results depend on both the type of recognizer and of detector. For organic video (see Figure 5), the ML recognizer works properly for faces detected through ML Detector, as shown by *Accuracy* and *Error* values of 95% and 5%, respectively. Yet, the same recognizer leads to poor results (*Accuracy* = 63%, *Error* = 37%) when applied to DL-detected face areas. The DL face recognizer considered in our study led to the same good performances for both ML and DL detected face areas: *Accuracy* = 99%, *Error* = 1%. The face recognition results obtained out of processing the second database are presented in Figure 6: note that as this database is quite small, a direct comparison among the values presented in Figure 5 and Figure 6 might be unfair.

Table 1 shows the end-to-end performances for the four working configurations and for the two databases. The results show that the organic video content comes across with particular constraints: the best configuration (the MMOD coupled to the Pruned ResNet model) affords an *Accuracy* = 85%, that is 5% lower than in the case of celebrities database. Note that while the *Precision* values are quite always equal to their ideal limit, the *Recall* is quite low: *Recall* = 0.8 for the same DL-DL and organic video database.

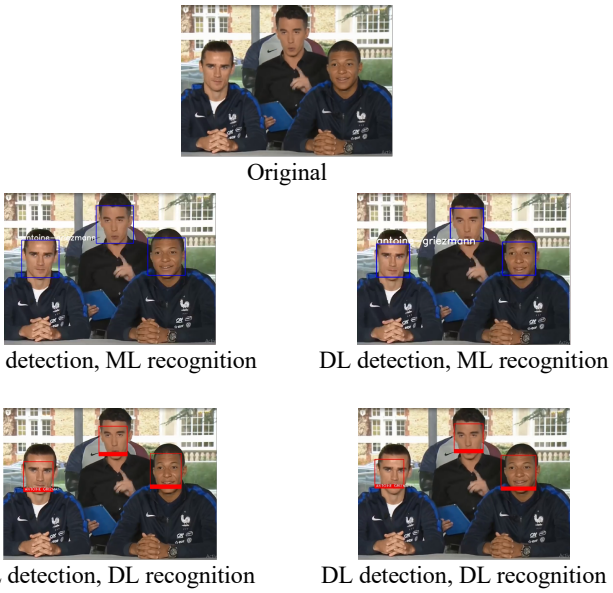

Figure 3. Illustration of the experimental results for frontal face images.

Figure 4. Illustration of the experimental results for non-frontal face images.

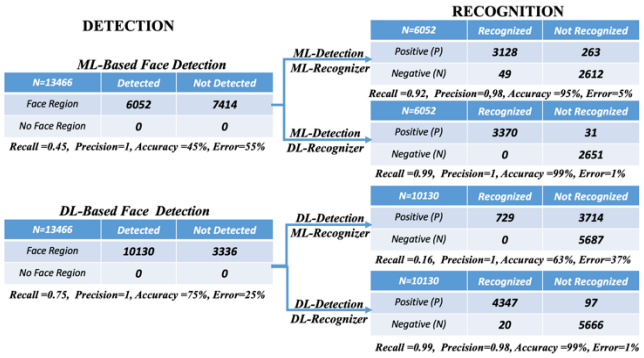

Figure 5. Results first database (organic video database)

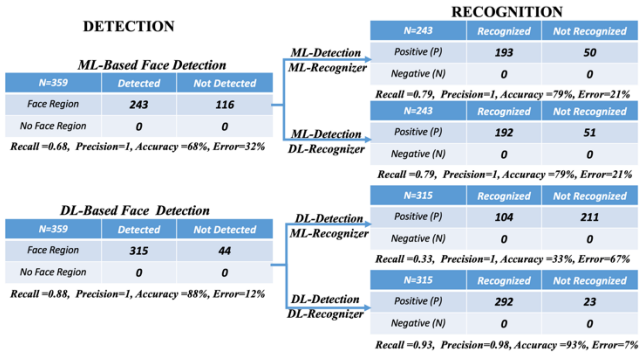

Figure 6. Results of the second database (Face Recognition Dataset)

Table 1. End-to-end performances for the four types of methods and for the two datsets.

	Method	Recall	Precision	Accuracy (%)	Error (%)
1st dataset	ML -ML	0.54	0.99	60	40
	ML - DL	0.56	1	62	38
	DL - ML	0.6	1	70	30
	DL - DL	0.8	0.99	85	15
2nd dataset	ML -ML	0.72	1	72	28
	ML - DL	0.72	1	72	28
	DL - ML	0.62	1	62	38
	DL - DL	0.9	1	90	10

4. CONCLUSION

The present paper reports on an experimental study on the usage of state-of-the-art face detection and recognition algorithm for the applicative framework of organic video, a particular video processing field characterized by completely unconstrained content recording and distribution conditions. Subjective, applicative-oriented evaluations (as illustrated in Figure 3 and Figure 4), as well as objective evaluations (as presented in Figure 5 and Figure 6, and in Table 1) bring to light that MMOD coupled to a Pruned ResNet model would practically outperform the other three combinations for organic video, resulting in an *Accuracy* = 0.85. This result is connected to the lack of sensitivity in the overall process: *Recall* = 0.8 while *Precision* = 0.99. Within this study, an organic video database has been collected and current efforts are devoted to making it publicly available. Future work will be devoted to extending the face detection and recognition task to more complex, multi-semantic / multi-contextual information detection (e.g. recognize a soccer player wearing its club shirt).

References

[1] G. Shakhnarovich, P. A. Viola, B. Moghaddam. "A Unified Learning Framework for Real Time Face Detection and Classification," Proceedings of the Fifth IEEE International Conference on Automatic Face and Gesture Recognition

[2] H. Gang, Y. Ming-Hsuan, L.M. Erik, M. Yi, T. Matthew, J.K. David, S.H. Thomas. "Introduction to the Special Section on Real-World Face Recognition"

[3] M. H. Yang, D. J. Kriegman, N. Ahuja. "Detecting Faces in Images: A Survey," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 24, no. 1 January 2002

[4] A. G. Raval, H. B. Bhadka. "A Summary of literature review on Face Recognition, " 2019 JETIR June 2019, Volume 6, Issue 6 (ISSN-2349-5162)

[5] S. Devadethan, G. Titus, S. Purushothaman. " Face detection and facial feature extraction based on a fusion of knowledge-based method and morphological image processing," 2014 Annual International Conference on Emerging Research Areas: Magnetics, Machines and Drives (AICERA/iCMMD)

[6] Liawa, G. You, F. Syu, C. Huang. "A New Knowledge-Based Face Image Indexing System through the Internet," ISBN 89-5519-129-4, Feb. 20-22, 2006 ICA0T2006

[7] L. Zhang, P. Lenders. "Knowledge-Based Eye Detection for Human Face Recognition," Fourth International conference Knowledge-based Intelligent Engheehg Systems 6 AIM Teclu News, 3@ Aug-I' Sept 2o00, Bingham. UK

[8] J. Zhao, Y. Cheng, Y. Xu, L. Xiong, J. Li, F. Zhao, K. Jayashree, S. Pranata, S. Shen, J. Xing, S. Yan, J. Feng. " Towards Pose Invariant Face Recognition in the Wild," 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition

[9] N. Liu, H. Wang. "Invariant Facial Features Under Pose Variations for Face Recognition," 0-7803-9282-5/05/2000, IEEE, ICICS 2005

[10] C. Rahmad, R. M. Asmara; D. R. H. Putra, I. Dharma, H. Darmono, I. Muhiqqin. "Comparison of Viola-Jones Haar Cascade Classifier and Histogram of Oriented Gradients (HOG) for face detection," IOP Conference Series: Materials Science and Engineering

[11] S. V. Viraktamath, M. Katti, A. Khatawkar, P. Kulkarni. "Face Detection and Tracking using OpenCV," The SIJ Transactions on Computer Networks & Communication Engineering (CNCE), Vol. 1, No. 3, July-August 2013

[12] A. Zarkasi, S. Nurmaini, D. Stiawan, Firdaus, H. Ubaya, Y. Sanjaya, Y. N. Kunang." Face Movement Detection Using Template Matching," International Conference On Electrical Engineering And Computer Science (ICECOS) 2018

[13] M. H. Teja. " Real-time Live Face Detection using Face Template Matching and DCT Energy Analysis," 2011 International Conference of Soft Computing and Pattern Recognition (SoCPaR)

[14] C. Molder, R. Oancea. "Appearance-based Facial Detection for Recognition," 2012 9th International Conference on Communications (COMM)

[15] A. Memiş, F. Karabiber. "Face recognition on mobile environment images using appearance-based methods," 2016 24th Signal Processing and Communication Application Conference (SIU)

[16] O. Saman, L. Stanciu, "Image Processing Algorithm for Appearance-Based Gesture Recognition," 2019 23rd International Conference on System Theory, Control and Computing (ICSTCC)

[17] N. Boyko, O. Basystiuk, N. Shakhovska. "Performance Evaluation and Comparison of Software for Face Recognition, based on Dlib and Opencv Library," IEEE Second International Conference on Data Stream Mining & Processing August 21-25, 2018, Lviv, Ukraine

[18] H. Rowley, S. Baluja, T. Kanade. "Neural Network-Based face detection," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 20, no. 1, pp. 23-38, Jan. 1998.

[19] K. Bryan. "Face Detection for Crowd Analysis Using Deep Convolutional Neural Networks," E. Pimenidis and C. Jayne (Eds.): EANN 2018

[20] M. Rezaei, H. Z. Nafchi, S. Morales. "Global Haar-like Features:A New Extension of Classic Haar Features for Efficient Face Detection in Noisy Images," Conference: 6th Pacific-Rim Symposium on Image and Video Technology, 2013

[21] E. K. Davis. "Max-Margin Object Detection," arXiv:1502.00046v1 [cs.CV] 31 Jan 2015

[22] S. Zhang, M. Turk. " Face recognition using eigenfaces," 1991 IEEE Computer Society Conference on Computer Vision and Pattern Recognition

[23] Zhujie, Y.L. Yu. "Face Recognition with Eigenfaces," Proceedings of 1994 IEEE International Conference on Industrial Technology - ICIT '94.

[24] M. Çarıkçı, F. Özen."A Face Recognition System Based on Eigenfaces Method," Procedia Technology Volume 1, 2012, Pages 118-123.

[25] M. U. Rahman. "A comparative study on face recognition techniques and neural network,"

[26] X. He, S. Yan; Y. Hu; P. Niyogi, H. J. Zhang; IEEE Transactions on Pattern Analysis and Machine Intelligence, pp. 328 – 340, 2005

[27] H. Lu, K. N. Plataniotis, A. N. V. Mpca. "Multilinear principal component analysis of tensor objects," IEEE Trans. on Neural Networks, 19(1):18–39, 2008.

[28] X. Zhı-Hua, L. Guo-Dong, F. Zhı-Jun. "A Novel Infrarerd Face Recognition Based On Local Binary Pattern," Proceedings of the 2011 International Conference on Wavelet Analysis and Pattern Recognition, Guilin, 10-13 July, 2011

[29] Y.K. Park, J.K. Kim. "Fast adaptive smoothing based on LBP for robust face recognition," Electronics Letters 22nd November 2007 Vol. 43 No. 24

[30] J. Krizaj, V. Struc, N. Pavesic. "Adaptation of SIFT Features for Robust Face Recognition," Faculty of Electrical Engineering, University of Ljubljana,Trzaska 25, SI-1000 Ljubljana, Slovenia

[31] H. Bay, A. Ess, T. Tuytelaars, L. Van Gool. "Speeded-Up Robust Features (SURF)," Computer Vision and Image Understanding, 110 (2008), pp. 346–359

[32] M. U. Rahman. "A comparative study on face recognition techniques and neural network,"

[33] X. Fan, B. Verma. "A comparative experimental analysis of separate and combined facial features for GA-ANN based technique," Sixth International Conference on Computational Intelligence and Multimedia Applications, 2005.

[34] S. Pang, D. Kim, S. Y. Bang. "Face membership authentication using SVM classification tree generated by membership-based LLE data partition," IEEE Transactions on Neural Networks, Volume: 16 , Issue: 2 , pp. 436 – 446, 2005

[35] K. Lu, X. He, J. Zhao. "Semi-supervised Support Vector Learning for Face Recognition," Lecture Notes in Computer Science, pp. 104-109, 2006

[36] N. Jamil, S. Iqbal, N. Iqbal. "Face recognition using neural networks," Technology for the 21st Century, pp. 277 – 281, IEEE INMIC, 2001

[37] P. Latha, L. Ganesan, S. Annadurai. "Face Recognition using Neural Networks," Signal Processing: An International Journal (SPIJ) Volume (3) : Issue (5). 153, 20091-)

[38] D. Farah, A. Aftab, A.D. Fayaz, M. Hira, G. Abddul, D. Deeba. "LBPH-based Enhanced Real-Time Face Recognition," (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 10, No. 5, 2019

[39] S. Chanda, A. Chakrapani, A. Brun, A. Hast, U. Pal, D. Doermann. "Face Recognition - A One-Shot Learning Perspective," 2019 15th International Conference on Signal-Image Technology & Internet-Based Systems (SITIS)

[40] G. Mingyu, C. Jianfeng, M. Hongbo, Q. Dawei. "A Transfer Residual Neural Network Based on ResNet-34 for Detection of Wood Knot Defects," College of Science, Northeast Forestry University, Harbin

[41] https://www.kaggle.com/vasukipatel/face-recognition-dataset

[42] L. Wandzik, G. Kaeding, R. V. Garci. "Morphing Detection Using a General-Purpose Face Recognition System," 2018 EUSIPCO

Author Biography

Yigit Oguzhan AKBAY holds an engineering degree from University of Zonguldak Bülent Ecevit in Turkey (2020) and an MS degree from University Paris-Saclay iny France (2021). The results presented in this paper were obtained during his MS internship at Telecom SudParis, in collaboration with Vidimizer (https://vidmizer.com).

Mihai Mitrea holds an HDR degree Pierre and Marie Curie University in France (2010) and a PhD from Politehnica University of Bucharest in Romania (2003). He is currently Associate Professor at Telecom SudParis. He is vice-president of the Cap Digital's Technical Commission on Digital Content and serves as an advisor for the French delegation at ISO/IEC JTC1 SC29 (a.k.a. MPEG