

An Efficient System for Partial Occluded Face Recognition

Haixiang Zhang¹(✉), Peng An¹, and Dexin Zhang²

¹ School of Electronic and Information Engineering,
Ningbo University of Technology, 201 Fenghua Road, Ningbo 315211, Zhejiang, China
sean_public@qq.com

² Tianjin ISecure Technologies Co. Ltd., TEDA MSD-G1 10F, TEDA Tianjin
Economic-Technological Development Area, Tianjin, China
<http://www.nbut.edu.cn>

Abstract. Face recognition has attracted a lot of interest and made a wide range of applications in the real world. However the technology's efficiency is limited to well positioned, clear face images, which is not always the case in reality. In this paper, we present a software system that is resilient to various image quality degradation, such as face occlusions, illumination effects and face postures. The system exploits deep learning to build an effective model for face detection, and an Elastic Graph Matching based method to extract the key face features for comparison with the face library. The system also provides flexible functionality and extensions to adopt other biometric recognition algorithm for cross-checking and provides a higher reliability. Experiments show that the proposed system offers a good performance and is robust to face occlusions.

Keywords: Biometrics · Face recognition · Partial occlusion
Deep learning · FPCA · EGM

1 Introduction

Computer vision and image processing advances in theories and techniques boosted development of image based applications. Benefiting from it, biometric recognition technologies have been developing quickly in recent years, and face recognition has seen arguably the most wide range of applications among all image analysis and understanding technologies [1, 2]. Although these machine recognition systems have brought more and more convenience and efficiency to our every day life, they still suffer from their limitations in many real application scenarios. For example, images acquired via surveillance cameras, which are very pervasive today, are often taken from an outdoor environment with variations of

H. Zhang—This work was partially supported by Chinese National Natural Science Fund 61502256.

illumination conditions; cameras installed at public passages, or any uncontrolled scenarios. Such images could record faces of different poses, and faces could be occluded by glasses, hat, scarf, or even just by other people. Performing face recognition on such image signals remains a challenging problem. Under such circumstances, certain part of the face information is occluded, making it hard to extract all the important features for recognition, and resulting in degraded recognition accuracy. A lot of work has been carried out to tackle these problems [3–8]. These works are mostly theoretical research tested with experiments in labs, but rarely adopted for practical applications. In this paper, we present a face recognition software system ready for practical usage and applications, which provides robust recognition accuracy with occluded face images, while maintaining a high flexibility for users to set parameter for retrieving purpose, so that it is possible to search for certain interested individuals among the massive amount of recorded videos. It can also be easily extended to a system that combines more than one biometric recognition algorithms where cross-checking helps to satisfy a higher security requirement.

2 The Challenges for a Robust Face Recognition System

Face recognition technology is based on human facial features. A typical face recognition system works in the following procedure: 1. The system is given an input of face image, or video stream. It first analyzes the content to determine whether there exists a human face in the scene. 2. If there is a face/s, it then further locates each face, detects the size and positions of the major facial organs of each face. 3. Based on the above information, the system then further extracts each individual’s unique identity characteristics. 4. The system compares with the known faces, normally from a database, to verify each person’s face identity. Above is a straightforward and sensible process. But in many cases, the collected face images are not standard front face images [4,5]. And it is quite often the case that a collected face image is partly occluded by glasses, light beard, or simply distorted by reflection and other lighting effects. Such image quality would heavily degrade the system’s accuracy of face recognition performance. Targeting at the above problems, the proposed system exploits deep learning algorithm to perform partial face modeling in the process and substantially improves the accuracy of face recognition. The proposed system can process the full clear recorded faces, as well as those occluded faces with glasses, beard and other obstructions, which helps to improve the searching accuracy and efficiency. In addition, the system can be used to provide an extra verification mechanism according to the needs of the security level. Cross-checking the identity using more than one biometric mechanism enables a higher level of security than the simple face recognition system [3]. The system can adapt to different security levels, with multi-mode biometric identification, such as Iris and Face recognition, to ensure an even higher security with a very brief processing time.

3 System Structure and the Sub-modules of the System

Figure 1 shows the structure of the proposed face-recognition system. As depicted in the figure, the system is organized as a series of layers. The lowest layer takes in the video or image signals, the layers in the middle level performs the face recognition function, and the highest application layer serves as an interface to support the actual security checking or individual searching job.

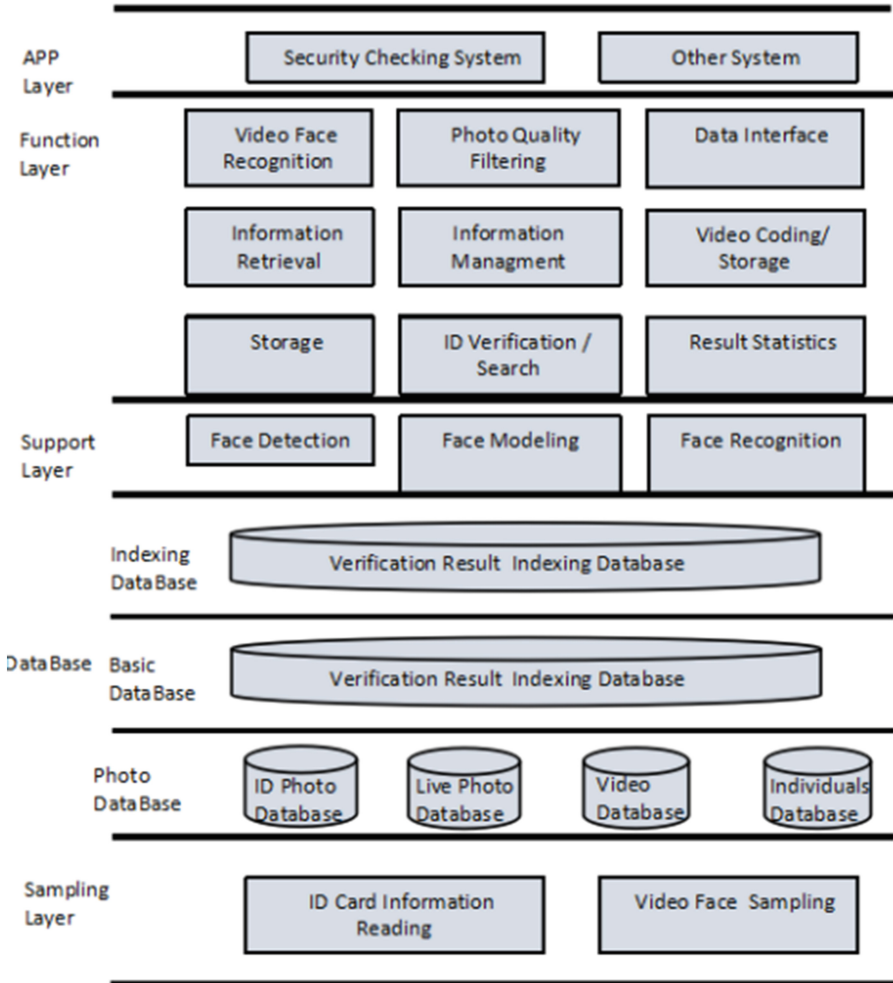


Fig. 1. System structure.

The sub-modules are described in details below in this section.

3.1 Face Detection and Locating

Here face detection refers to the process of judging whether there are human faces in an image with dynamic scene or a complex background. If there are faces, they would be then extracted. Face detection is the basic stage of face recognition, and only with the presence of a face can the recognition process be carried out. Most of the face detection techniques in the current market only works on clear and complete face images. The proposed system integrates a set of face features obtained from a large number of partial occluded face images using deep learning methods. It can process the incomplete faces and thus improve the performance under complex scenarios.

Factors Affecting Face Locating. Locating the faces in an image is not always straightforward. To perform face detection and locating, the following factors need to be considered: 1. the position, angle and the posture of the individual in the image; 2. The variable size of the face region in an image; 3. The illumination effects. In some situations the photo shooting is strictly under control and the image quality is usually guaranteed. For example, when the police takes photos of a detained suspect, the face is always aligned to the frame and locating the face in the image is simple. A typical ID photo also has very simple background color and standardized face presentation, therefore locating a face is also easy in this case. However, in many other real life cases, human faces could be recorded without regular positioning, such as video recorded by surveillance camera, or with highly dynamic background changes, such as video recorded by hand held mobile phone. Thus the faces' positions are unpredictable with these complex conditions.

Models for Face Detection. Contours and skin tones are important characters of a face. They are generally stable and enable a face to stand out of most background objects, hence, they can be used as good candidates for fast face detection in color images. The essential procedure of the feature detection method is, first, developing the skin color model, and use it to detect the pixels of skin color, then locate the possible face area according to the similarity and correlation in spatial and chromatic domains.

The picked possible face areas are then fed into the face model developed through machine learning method to go through further checking before the ultimate judgment of whether the detected area contains a face. In the process of building the database, we have included a large number of partial occluded face samples, including glasses, or beard occlusion, and the output model works effectively on such faces. Some occluded face detection results are shown in Fig. 2.

3.2 Face Alignment Module

Once the face detection is finished, the picture must go through some pre-processing stage, which is essential to keep the accuracy of the face recognition,

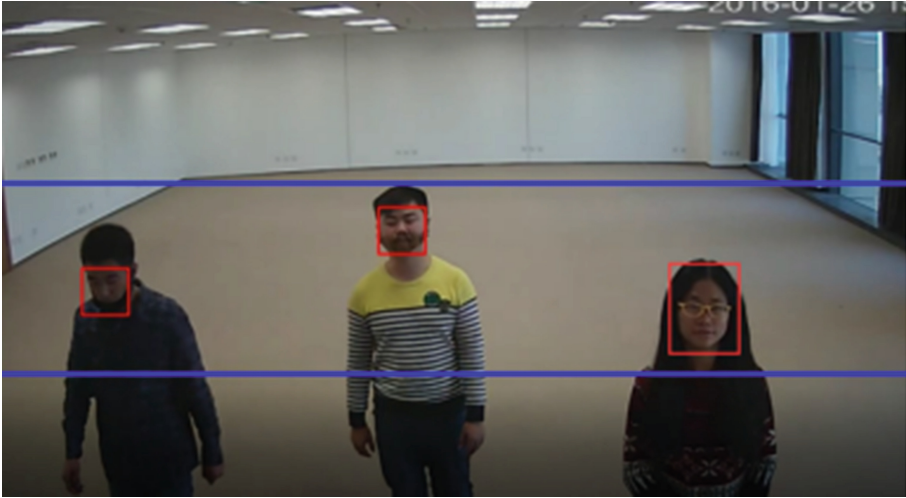


Fig. 2. Detection of occluded faces.

especially in a complex environment. This stage performs size and gray scale normalization, head posture correction, image segmentation etc. The purpose of performing these operations is to improve the image quality, such as reducing noise, achieving uniform image gray scale and size, which sets a good condition for feature extraction and classified identification in a later stage. Face alignment has two parts, geometric normalization and gray scale color normalization.

Geometric Normalization. Geometric normalization involves two steps: face correction and face cropping. The detected expression sub-image is transformed into a uniform size, which is conducive to the extraction of facial features.

Grayscale Normalization. The gray scale normalization is mainly to increase the contrast of the image and illumination compensation. It aims at increasing the brightness of the image, so that the details of the image are clearer, and the impact of light effects are reduced. Some resultant images of face alignment faces are shown below. Though the photos are originally taken at different postures and angles, it is clear that they are now of uniform size and presentation, and the eyes, nose mouth etc. are at similar positions of the picture, which makes it easier to extract face features from the faces. Such visually similar images are ready for further processing and comparison (Fig. 3).

3.3 Partial Occluded Face Modeling

Partial occlusion of the face often results in loss of important information of the face, thus affecting the face recognition accuracy. The proposed Partial Occluded



Fig. 3. The resultant face images after alignment.

Face Modeling software uses a fuzzy principal component analysis (FPCA) to carry out a separate modeling of occluded faces [9, 10]. The process is as following: first, an occluded face is projected onto the eigenface space and a reconstructed face is then obtained by a linear combination of eigenfaces. The difference between the reconstructed face image and the original image is calculated, and then passed through weighted filter to calculate a probability value of the face part being occluded. This value is then used as a coefficient to combine the original image and reconstructed image to form a new face image. In the subsequent iterations, this coefficient is used in the FPCA for reconstruction, and the

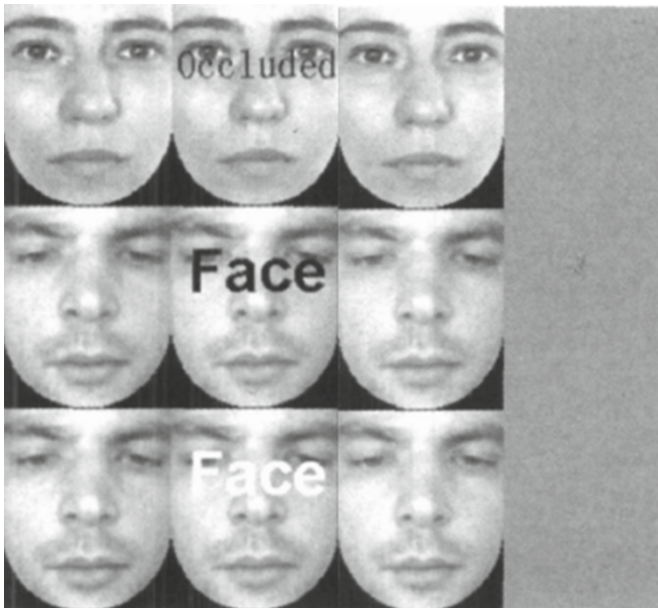


Fig. 4. Reconstruction of occluded face images.

cumulative error is used for occlusion detection. This approach can accurately locate the face occlusion area, and get a smooth and natural reconstruction of the face image. Some partially occluded faces and the corresponding reconstructed face image are shown below in Fig. 4.

The images are shown in columns side by side for comparison purpose. The first column are the original clear face images, the second column are the occluded face images, while the third column is FPCA method reconstruction of the face images. The fourth column shows the differences between the reconstructed faces and the original faces. It can be seen from the figure that the FPCA method could recover the occluded part of the face and reconstruct a natural and smooth face image. These reconstructed images are very close to the original ones, hence it can be expected that feature extraction over these images would be close to originals.

3.4 Key Face Feature Extraction

Feature extraction is the process that tries to find the clear, stable and effective face information with the presence of interference and noise in the image. The facial feature extraction method can also analyze the environment to extract the facial features with different algorithm. Face feature extraction is the core step in face recognition, which directly determines the recognition accuracy. The proposed software uses an Elastic Graph Matching method, which is based on dynamic link architecture (DLA) [11–14]. This method creates a property map for the face in a two-dimensional space, and places a topology map over the human face. Each node of the map contains an eigenvector, which records the distribution information of the face near the node [15, 16]. The topological connections between nodes are denoted with geometric distance, thus forming a two-dimensional topology description of the face. When performing the face recognition with this system, we can simultaneously consider both node eigenvector matching and relative geometric position matching. In one way, we can scan the topology map structure on the face image to extract the corresponding node feature vector, and use the distance of different positions between the topological maps of the image and the face pattern in the library as the similarity measurement. Additionally, an energy function can be used to evaluate the matching between the target face vector field and the known face vector field, or, the minimum energy function matching. The method is robust to illumination and posture changes. The main drawback of this method is the high computational complexity, since calculating the model map must be performed over every individual face image, which takes up a lot of memory (Fig. 5).

In experiments over databases, the method showed a decent performance, and it adapts well to face posture and facial expression changes.

3.5 Face Recognition Software

Face recognition software works with all the above sub-modules combined. When fed in the photo of a target person, the system automatically detects and locates

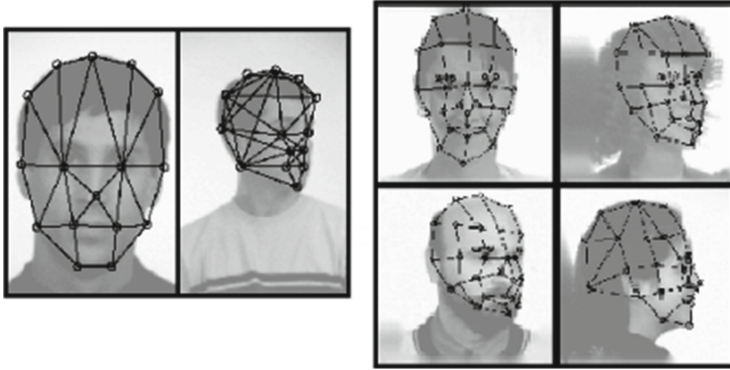


Fig. 5. Generalized elastic graph matching for face location and recognition

the face within the photo, then extracts the face image, process the image and calculates the key features of the face, then compare it with the features of recorded faces in the database. If matches are found, the detected faces are then highlighted to indicate positive match. Figure 6 shows a case of bearded face recognition, which is a typical case in practice. The individual was wearing a quite long beard and the system still accurately picks him out, and highlighted the face with a red square, as shown below.



Fig. 6. Detection of bearded face

Thanks to the earlier stages of the system, which has already reconstructed a smooth and natural face image, removed the negative effect of occlusion and other image degradation, the system recognition accuracy can be still maintained, making the system robust to such interference. In practice, the system could also enable the users to set a series of parameters, such as similarity, time of recording, and starts to search through the library. Finally, the found similar faces are displayed, and a quick confirmation by human checking is made possible.

3.6 Robustness of the Recognition

The proposed system is implemented and installed at various venues and public passages for testing. Volunteers have been walking before the cameras and the recorded videos are processed. Some results of typical occlusions are tested in these experiments, and the results are shown below in Table 1.

Table 1. Recognition Rate with Different occlusions.

Glasses	Beard	Scarves
97.13%	98.39%	99.32%

It is clearly shown that the tested typical occlusions can be effectively tackled and it achieves a good recognition rate, close to clear face images in many cases. Other tests performed include illumination affected images, and different face postures, where the system also showed substantial effectiveness. The results are not included here as it is hard to define such conditions. This part will be presented in a separate publication. In general, the system demonstrates a very good recognition performance in real life scenarios.

4 Face Recognition for Person Retrieval

Retrieving certain target person or objects is a frequent and highly desired job in practical application of video systems. Close circuit TV (CCTV) systems have shown their fastest growth in the past 2 decades. All kinds of video cameras have been installed, almost everywhere, and the number is still growing exponentially. Such system monitors a fixed area and tries to maintain a seamless records of video coverage of that area. The wide usage of such systems provides a big resource for event and human tracing. They have found many useful cases in practice, especially they provided countless examples to helping police and other public security departments to solve crime cases. However, these systems are designed to record the video signals while little has been considered about how to process these recorded files. Many systems in applications were installed in quite early years, which left a lot of room for efficiency and accuracy improvement. Some common problems have been widely seen, such as, Humans are

still the major searching tools used for any specific information retrieving jobs. The CCTV systems only record the videos in case any of them could include some key information, but it is still up to manpower to search for such information. Typically, the video reviewers have to sit before the monitors for very long hours. Even when there is no active objects appearing in the scene, they still need to go through every second of the video without accelerating. Such working procedure not only result in an inefficient and low quality searching, but also harms the physical and mental health of the reviewers. Apart from the massive amount of accumulated videos, the video quality forms another challenge to make use of the files. CCTVs in real life are installed everywhere, over a long period, so the quality of recorded video varies dramatically. This could be due to different cameras, or most often, due to people's different postures when being recorded. Real life camera recording is a typical uncontrolled scenario and the resultant videos could suffer from all kinds of quality problems, such as occlusion, poor shooting angle or illumination. All these make processing the files difficult. With robust face recognition, it is then possible to search through these files effectively. The computers can save humans from manually browsing the huge amount of videos, while the algorithm ensures an effective search among the imperfect videos. Figure 7 shows the interface window of a face-recognition based people searching function in the proposed system. It can be seen in the figure that the people in the image are either of poor shooting angle, or have the face partially occluded, yet the system could still find the targeted individuals from the video. The algorithm to recognize the occluded faces advances the technology to a much larger range of applications, especially for police and public security departments, where browsing and searching through



Fig. 7. Face recognition searching module.

video are mostly performed by humans manually, which is not only tedious but also inaccurate. The proposed system provides a perfect alternative approach.

5 Conclusion

Biometrics technologies, especially face recognition, has attracted a lot of interest in both industry and academic field, When applied in practice, the algorithms often suffer from occlusion and other interference. In this paper, we propose a robust face recognition system, which tackles such interference. The system makes use of face model to reconstruct the occluded area and produces a smooth and natural face for further processing. Key features of the face can then be extracted, and the recognition accuracy is improved to be close to non-occluded faces. This robust recognition algorithm then enables effective searching through the massive amount of videos recorded by CCTV systems where video quality is not guaranteed. The system provides a searching function over video files, and it shows a reliable and efficient performance when applied in practice.

References

1. Zhao, W., et al.: Face recognition: a literature survey. *ACM Comput. Surv.* **35**(4), 399–458 (2003)
2. Ahonen, T., Hadid, A., Pietikinen, M.: Face description with local binary patterns: application to face recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* **28**(12), 2037 (2006)
3. Zhang, H.: A multi-model biometric image acquisition system. *Biometric Recognition. LNCS*, vol. 9428, pp. 516–525. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-25417-3_61
4. Tarres, F., Rama, A.: A novel method for face recognition under partial occlusion or facial expression variations. In: *IEEE International Symposium on ELMAR 2005*, pp. 163–166 (2005)
5. Andrés, A.M., et al.: Face recognition on partially occluded images using compressed sensing. *Pattern Recogn. Lett.* **36**(1), 235–242 (2014)
6. Zhang, X., Gao, Y.: Face recognition across pose: a review. *Pattern Recogn.* **42**(11), 2876–2896 (2009)
7. Ding, C., Tao, D.: Trunk-branch ensemble convolutional neural networks for video-based face recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* **99**, 1 (2017)
8. Yang, A.Y, et al.: Fast 1-minimization algorithms and an application in robust face recognition: a review. In: *IEEE International Conference on Image Processing*, pp. 1849–1852. IEEE (2010)
9. Sharma, S.: *Applied Multivariate Techniques*, pp. 100–101. Wiley (1996)
10. Partridge, M., Jabri, M.: Robust principal component analysis. *J. ACM (JACM)* **58**(3), 1–73 (2011)
11. Wiskott, L., et al.: Face recognition by elastic bunch graph matching. *IEEE Trans. Pattern Anal. Mach. Intell.* **19**(7), 775–779 (1997)
12. Shin, S., Kim, S.-D., Choi, H.-C.: Generalized elastic graph matching for face recognition. *Pattern Recogn. Lett.* **28**(9), 1077–1082 (2007)

13. Wiskott, L., Fellous, J.M., Kruger, N., Malsburg, C.V.D.: Face recognition by elastic bunch graph matching. *IEEE Trans. Pattern Anal. Mach. Intell.* **19**(7), 775–779 (1997)
14. Lades, M., Vorbruggen, J.C., Buhmann, J., Lange, J., Malsburg, C.V.D., Wurtz, R.P., et al.: Distortion invariant object recognition in the dynamic link architecture. *IEEE Trans. Comput.* **42**(3), 300–311 (1993)
15. Yambor, W.S.: Analyzing PCA-based face recognition algorithms: eigenvector selection and distance measures, pp. 39–60. All Publications (2002)
16. Han, P.Y., Jin, A.T.B., Siong, L.H.: Eigenvector weighting function in face recognition. *Discrete Dyn. Nat. Soc.* 2011(2011–02-16), 2011(1026–0226), 701–716 (2011)