Application of robust face recognition in video surveillance systems*

ZHANG De-xin (张德馨)[1], **AN Peng** (安鹏)[2], **and ZHANG Hao-xiang** (张浩向)[2]**

1. Tianjin ISecure Technologies Co. Ltd, Tianjin 300457, China

2. School of Electronic and Information Engineering, Ningbo University of Technology, Ningbo 315211, China

(Received 31 August 2017; Revised 23 November 2017)

In this paper, we propose a video searching system that utilizes face recognition as searching indexing feature. As the applications of video cameras have great increase in recent years, face recognition makes a perfect fit for searching targeted individuals within the vast amount of video data. However, the performance of such searching depends on the quality of face images recorded in the video signals. Since the surveillance video cameras record videos without fixed postures for the object, face occlusion is very common in everyday video. The proposed system builds a model for occluded faces using fuzzy principal component analysis (FPCA), and reconstructs the human faces with the available information. Experimental results show that the system has very high efficiency in processing the real life videos, and it is very robust to various kinds of face occlusions. Hence it can relieve people reviewers from the front of the monitors and greatly enhances the efficiency as well. The proposed system has been installed and applied in various environments and has already demonstrated its power by helping solving real cases.

Document code: A **Article ID:** 1673-1905(2018)02-0152-4

DOI https://doi.org/10.1007/s11801-018-7199-6

Theoretical advances and technical progress have helped to boost the development of biometrics in recent years[1-4]. Biometric technologies exploit the biological characteristics of humans, such as face, fingerprint, iris, vein, and gait to recognize identities. A lot of research work has been done to enable computers to understand and extract these biological characteristics.

Biometric technologies are applied mainly for two purposes, identity verification and searching for targeted persons within database. The former has been found useful in passage and access control, attendance checking, and payment authorization, while the latter is now exploited for processing a large amount of data, typically video files, where manual checking is far too exhausting, slow, very inefficient and inaccurate. With the wide application of monitoring video cameras, or close circuit TV (CCTV) systems, the video signals increase exponentially, which provides more and more application scenarios for biometrics as a searching algorithm.

In this paper, we propose a system that exploits biometric techniques to process massive video files. In any case, the recorded video signals need to be checked through for any specific individual person or event, and the proposed system could automatically carry out the whole process. Especially, the system can effectively process the occluded face images, hence demonstrates very high accuracy.

The wide application of video cameras means a huge amount of video data are produced everyday. Close circuit TV (CCTV) systems have shown their fastest growth in the past two decades, and the number is still growing exponentially. Such a surveillance system monitors a fixed area and tries to maintain a seamless record of video coverage of that area. The usage of such systems provides a big resource for event and individual person tracking. They have found many useful cases in practice, especially they provided countless examples to help police to solve crime cases.

Automatic algorithms have to be adopted in order to make efficient use of the video data. Most of the systems are designed to record the video signals while few have considered how to process these recorded data. One of the most concerning challenges they face is the lack of efficient searching mechanisms. In practice, it is still mostly up to manpower to perform any individual person retrieving job. This not only means the video reviewers have to sit before the monitors for very long hours, but also means unreliable and inaccurate results. Biometric-based algorithms enable the computers to understand the person or event characteristics, so the video files can be processed by software instead of humans.

The face recognition algorithms provide a tireless and

* This work has been supported by the National Natural Science Foundation of China (No.61502256). This paper was presented in part at the CCF Chinese Conference on Computer Vision, Tianjin, 2017. This paper was recommended by the program committee.

** E-mail: sean_public@qq.com

uniform searching job, but the performance suffers from poor face image quality. In many cases, the collected face images are not standard front face images[4,5], and it is common that a collected face image is partly occluded by glasses, light beard, or simply distorted by lighting effects. So the image quality would heavily degrade the system's accuracy of face recognition[3-8]. Therefore, an ideal system should offer a face recognition algorithm that is robust to such facial image distortions, in order that efficient searching through surveillance videos for a specific person can be achieved.

In content retrieval among a huge amount of videos, only automatic and intelligent techniques are feasible solution[9,10]. The proposed system exploits deep learning algorithm to perform partial face modeling in the process and substantially improves the accuracy of face recognition with occluded face images.

The system consists of a series of layers. The lowest layer takes in the video or image signals, the layers in the middle level perform the face recognition function, and the highest application layer serves as an interface to support the actual security checking or individual searching job.

Here "face detection" refers to the process of judging whether there are human faces in an image. Face detection is the basic stage of face recognition. If there are faces been detected, they would be then extracted for further processing. 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.

Contours and skin tones are important characteristics 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 developing the skin color model, and using it to detect the pixels of skin color, then locating the possible face area according to the similarity and correlation in spatial and chromatic domains.

The extracted possible face areas are then fed into the face model developed through machine learning method for further checking, before the ultimate judgment is made. In the process of building the database, we have included a large number of partial occluded face samples, and the output model works effectively on such faces.

Some face detection results with common occlusions are shown in Fig.1.

Once the face detection is finished, the picture must go through some pre-processing stage. This stage performs size and gray scale normalization, head posture correction, and image segmentation, etc.

Face alignment has two parts, geometric normalization and gray scale color normalization. Geometric normalization involves two steps: face correction and face cropping. 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 is reduced.

Fig.1 Detection of occluded faces

Some resultant images of face alignment are shown below in Fig.2. It is clear that they are now of uniform size and presentation, and the eyes, nose and mouth are at similar positions of the picture, which makes it easier to extract features from the faces.

Fig.2 The resultant face images after alignment

Partial occlusion of the face often results in loss of important information of the face, thus affecting the face recognition accuracy. The proposed system uses a fuzzy principal component analysis (FPCA) to carry out a separate modeling of occluded faces[11,12].

The process is as follows: 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 passes 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 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 images are shown below in Fig.3.

These reconstructed images are very close to the original ones, hence it can be expected that features extracted from these images would be close to those obtained from the original ones.

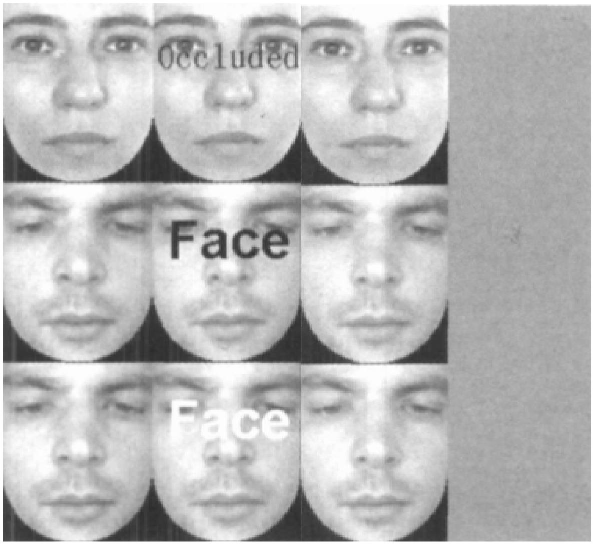

Fig.3 Reconstruction of occluded face images

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 proposed software uses an elastic graph matching method, which is based on dynamic link architecture (DLA)[13,14]. 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. 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.

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

Face recognition software works with all the above sub-modules combined. Fig.4 shows a case of bearded face recognition, which is a typical case in practice. The individual is wearing a quite long beard and the system still accurately picks him out, and highlights the face with a square, as shown below.

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 interferences.

Fig.4 Detection of beard face

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 by the system. Some results of typical occlusions are tested in these experiments, and the results are shown in Tab.1.

Tab.1 Recognition rates 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 shows 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.

With a robust face recognition algorithm, it is now 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. Fig.5 shows the interface window of a face-recognition based people searching function in the proposed system. It can be seen that the people in the image are either of poor shooting angle, or have the faces partially occluded, yet the system could still find the targeted individuals from the video.

Fig.5 The face recognition searching module

The algorithm to recognize the occluded faces advances the technology to a much larger range of applications. The proposed system provides a perfect alternative approach to manual browsing. In a typical case, the system can finish searching 20 h worth of videos within 1 h. Fig.6 is a screenshot when searching for a target person in the system.

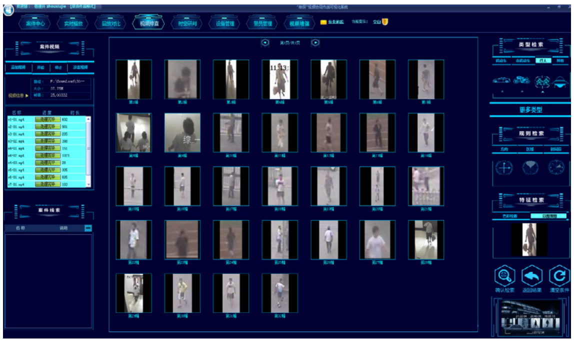

Fig.6 The result of suspect searching

Biometric technologies, especially face recognition, are very suitable for helping to make the normal CCTV monitoring system intelligent, so that the management and searching of the recorded videos could be performed by machines, thus greatly improving the efficiency and accuracy. We propose a system that takes these advantages of biometric technologies and provides a structured searching criterion setting function, and a fast, efficient, and accurate system is built. The robustness to poor face image conditions enables the system to cover most of the cases, where good and poor video qualities can both be processed. It has been installed and used in public security departments of various places. Applications in practical scenarios show that the system has outstanding performance, and traces down key information to help solving real cases.

References

[1] Zhao W, Chellappa R, Phillips PJ and Rosenfeld A, ACM Computing Surveys **35**, 399 (2003).

[2] Solanki K and Pittalia P, International Journal of Computer Applications **133**, 20 (2016).

[3] Zhang Haoxiang, A Multi-model Biometric Image Acquisition System, Biometric Recognition, Springer International Publishing, 2015.

[4] Cheng L, Wang J and Gong Y, Robust Deep Auto-encoder for Occluded Face Recognition, ACM International Conference on Multimedia, 1099 (2015).

[5] Padovani S., Tepper M. and Jacobo-Berlles J., Pattern Recognition Letters **36**, 235 (2014).

[6] Ou W, Luan X and Gou J, Robust Discriminative Nonnegative Dictionary Learning for Occluded Face Recognition, Pattern Recognition Letters, Available online 17 July 2017.

[7] Zhou B., Lapedriza A., Khosla A., Oliva A. and Torralba A., IEEE Transactions on Pattern Analysis & Machine Intelligence **PP**, 1-1 (2017).

[8] Yang A. Y, Fast ℓ1-minimization Algorithms and an Application in Robust Face Recognition: A Review, IEEE International Conference on Image Processing, 1849 (2010).

[9] Yang Y., Lovell B. C. and Dadgostar F., Content-Based Video Retrieval (CBVR) System for CCTV Surveillance Videos, Digital Image Computing: Techniques and Applications, 183 (2010).

[10] Sharma Subhash, Applied Multivariate Techniques, J. Wiley, 100 (1996).

[11] Partridge M. and Jabri M., Journal of the ACM **58**, 1 (2011).

[12] He E J, Fernandez J A and Kumar B V K V, Masked Correlation Filters for Partially Occluded Face Recognition, IEEE ICASSP, 1293 (2016).

[13] Wiskott L., IEEE Transactions on Pattern Analysis & Machine Intelligence **19**, 775 (1997).

[14] Hochul Shin, Seong-Dae Kim and Hae-Chul Choi, Pattern Recognition Letters **28**, 1077 (2007).

[15] Yambor W. S., Analyzing PCA-based Face Recognition Algorithms: Eigenvector Selection and Distance Measures, All Publications, 39 (2002).

[16] Han P. Y., Jin A. T. B. and Siong L. H., Discrete Dynamics in Nature and Society **2011**, 701 (2011).