

Facial Standardization Method in Video Face Recognition System

Kaiyu Wang^{1, a}, Zongmin Yu^{1, b}, Menglin Sheng^{1, c}, Peng Sun^{1, d}

¹Department of Electronic Science, Dalian University of Technology, Dalian, Liaoning, 116024, China

^aemail: wkaiyu@dlut.edu.cn, ^bemail: 554698162@qq.com

^cemail: 1928611293@qq.com, ^demail: sunpeng@dlut.edu.cn

Keywords: Face Recognition; Facial Standardization; Uncontrolled Environment

Abstract. In order to the practical face recognition system using in an uncontrolled environment exist many outside interference factors. In this paper, it is proposed that a novel facial standardization method to extract the core region of face samples, furthermore, simultaneously adjust the facial posture and achieve face alignment. It can avoid “mis-alignment disaster” during the feature matching process. In this paper, a video face recognition system of small samples capacity has been established to test this method. The experiments show that the method can make the face samples optimal in the uncontrolled environment and enhance the recognition rate of the face recognition system.

Introduction

With the development of face recognition technology, some commercial face recognition systems have been put into application. Because of the limitation of face recognition method, the majority of face recognition products can only be used by users cooperate, which greatly limits the scope of face recognition application. At present, face recognition in non-cooperation and uncontrolled environment cause the most attention of researchers[1-2]. In this environment, there are many posture changes[3], illumination changes[4] and expression changes[5] for the acquired face sample images.

In past time, the face recognition research works mostly focus on the effective realization of the algorithm, these algorithms are normally verify in the existing face standardization databases(such as ORL, YaleB, CMU-PIE), and have achieved very high recognition rate. However, the face samples in these databases have a fixed shooting environment, face region of uniform size and position and single hairdo. But the sample image obtained by face detection methods have different sizes, postures or positions and complex background. The face recognition algorithms under this harsh sample condition can not play maximum effectiveness, so that the recognition rate greatly reduced. For this reason, this paper proposes a novel facial standardization method, which can extract the core region of face, adjust postures, align the face features and achieve illumination and size normalization, to avoid “mis-alignment disaster” during the feature matching process.

Related work

About practical video face recognition algorithm, the researchers normally focus on illumination change and posture change. Compared with posture change, there are many effective solution[6] for illumination change questions, which can eliminate the effects of illumination change by illumination compensation, gradient descriptor and so on. About facial posture analysis, [7] proposes a facial reconstruction method based on deep learning and have made an excellent effect. But this method also need face alignment and large amount of calculation. [8] is based on probability and statistics thought to analyze sequential attitude changes of each person in the video. In the stage of recognition algorithm bottleneck, researches begin to pay more attention to how to exert the maximum capacity of existing algorithms in practical face recognition products. Face alignment is an important method to ensure the sample feature rectification. [9] use head pose estimation method to align facial features. [10] propose a novel cascaded regression framework for

face alignment based on a deep convolutional neural network. In addition, [11] propose a method of video-based face image quality assessment and filtration to pick out the high quality face samples for the recognition classifier in order to enhance the recognition rate.

In this paper, we combine various thoughts obtained from the above methods. Face standardization process is divided into five parts: landmark localization, geometric rectification, core region extraction, size normalization and illumination normalization. The paper establishes a simple practical video face recognition system using this standardization method as its preprocessing module, the system framework shown in Figure 1. The system has three modules: face detection module, preprocessing module and face recognition module. According to the characteristic of the face sample obtained from the detection module, the paper mainly introduce the face standardization process achieved by preprocessing module, and output the face samples for the recognition module to verify.

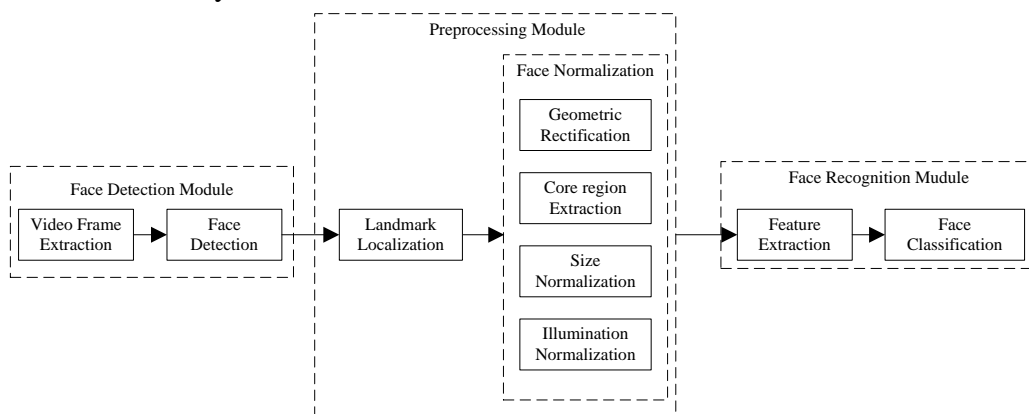


Fig.1. Video-based face recognition system frame chart

Landmark Localization

Active Shape Models(ASM)[12] is a facial features extraction method based on statistical learning model. As a classical method for landmark localization, ASM has been optimized constantly by researches. STASM[13] is an improved method based on ASM and has greater efficiency and accuracy. The STASM includes training process and matching process. In training process, MUCT face database(3755 face images, label 76 feature points artificially on each face) is used to train the models, and select the feature points satisfy the following characteristics: edge point, large curvature, T-connection points and equal diversion points of these points. These points will be ordered to constitute feature vector, then construct the shape statistical models after reducing dimensions by principal component analysis(PCA). In matching section, the face image got the most matching model by rotating, scaling, shifting the shape models trained before and adjust the parameters. As shown in Figure 2, the 77 feature points are labeled by STASM in a certain order.



Fig.2. The result of landmark localization[14]

Geometric Rectification

The head attitude changes are divided into three degrees of freedom: up and down, left and right, clockwise and anticlockwise. In this section, the head rotation changes are rectified based on geometric position of eyes. The coordinates of inner corners of two eyes $(x_1, y_1), (x_2, y_2)$ are extracted by the method section 3.1 mentioned. The rotation angle θ is calculated by Equation 1, the center point of two eyes (x_0, y_0) is considered as the axes to achieve face rotation correction, the process is shown in Figure 3. The coordinate of the feature points of the image after rotated can be calculated by Equation 2.

$$\theta = \arctan\left(\frac{y_2 - y_1}{x_2 - x_1}\right) \tag{1}$$

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} x_0 \\ y_0 \end{bmatrix} + \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} x - x_0 \\ y - y_0 \end{bmatrix} \tag{2}$$

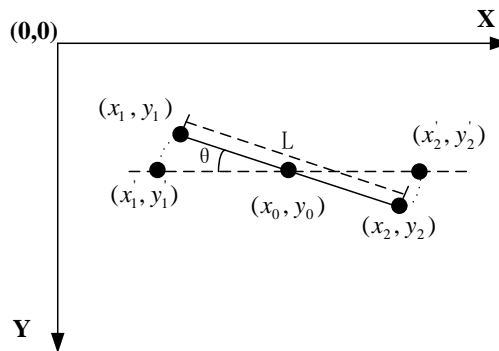


Fig.3. The schematic diagram of rotation correction

Core Region Extraction

From the research view of facial expressions, the lower half of the human face especially mouth is seriously affected by the expressions. Therefore, the paper analyzes the location information of nose, eyes and brows to confirm the core facial region.

A common face generally conforms to a geometric rule named “three courts and five eyes”. As shown in Figure 4, Face region in the vertical direction can be divided into five parts, and in the horizontal direction can be divided into three parts. The vertical equal division unit named d_H and the horizontal equal division unit named d_W can be calculated by Equation 3 and 4, y_n and y_e in Equation 3 are the vertical coordinates of nose tip and brow, and the x in Equation 4 are the horizontal coordinates of two inner eye corners. After a lot of attempts, we use the center point of two eyes (x_0, y_0) as datum point, then cut out $1.6d_W$ width region from each left and right side and $0.5d_H$ height region from upward side and $1.5d_H$ height region from downward side. But these parameters may not suitable when the facial expressions are extremely exaggerated, we can slightly modify the parameters according to the concrete situation.

$$d_H = y_n - y_e \tag{3}$$

$$d_W = x_2' - x_1' \tag{4}$$

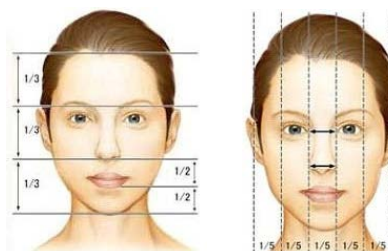


Fig.4. The schematic diagram of “three courts and five eyes” rule

Size Normalization

The feature extraction requires that all samples have the same dimension, so the size of samples must be unified. Such as the sample size of Yale face database is 100*100, CMU_PIE database has two sizes of 32*32 and 64*64, and ORL database is 92*112. Because of the sample core regions extracted by the method of section 3.3 are rectangle, if they are compressed into a square region, vertical direction of samples will shape change obviously. Therefore, we compress the samples into 92*112 uniformly like ORL database, and the result shows that facial feature maintains well.

Illumination Normalization

Illumination change is an important characteristic of uncontrolled environment. As the last link of preprocessing module after feature alignment and size normalization, the calculated quantity is lower and the effect is better than put it before others. The method named Gamma Intensity Correction(GIC)[15] are used in this section.

Suppose that the face sample of unknown illumination is I , and the face sample of illumination normalization is I_0 , I_{xy} is the gray value of point (x,y) . The equations are as follows:

$$I'_{xy} = G(I_{xy}; \gamma^*) \quad (5)$$

$$\gamma^* = \arg \min_{\gamma} \sum_{x,y} [G(I_{xy}; \gamma) - I_0(x,y)]^2 \quad (6)$$

$$G(I_{xy}; \gamma) = c \cdot I_{xy}^{\frac{1}{\gamma}} \quad (7)$$

Equation 7 is the Gamma transform; c is a gray stretch parameter, and γ is the Gamma coefficient.

Test platform Design

In order to test the facial standardization method the paper proposed, we establish a simple video-based face recognition system in Visual Studio 2010 based on Opencv. The face detection module of this system uses the adaboost cascade classifier with haar-like features[16] to find faces in video frame, and roughly extracts the face samples for the preprocessing module. The preprocessing module disposes these samples by 5 proposed steps, and output the standardization sample to the recognition module. Eigenfaces[17] and Fisherfaces[18] which are two classical face recognition algorithms are used to classify the samples.

Experiments

In this paper, all the test samples are derived from recorded video, and part of training samples are shown in Figure 5. These samples are extracted from video frame and disposed by preprocessing module. To ensure the accuracy of facial feature location, the deflection angle is maintained within $\pm 45^\circ$, and each person in the training database contains at least five different samples. The test system extracts 780 frames from the video, and the result is shown in Table 1.

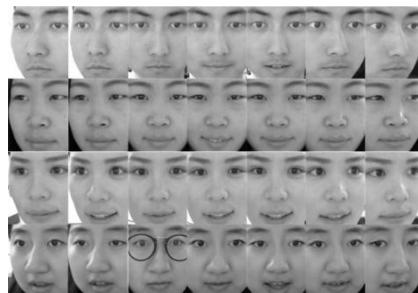


Fig.5. Part of training samples

Table 1. The recognition rate of the system

Method	Correct recognition (%)	False recognition (%)	Unable recognition (%)
Eigenfaces	59.4	36.4	4.2
Proposed+Eigen	74.5	17.9	7.6
Fisherfaces	71.3	24.5	4.2
Proposed+Fisher	88.3	4.1	7.6

From table 1, we can see that the proposed method is effective to improve the recognition rate of video-based system using classical recognition algorithms. But the proposed method also improve the unable recognition, the reason is that the attitude of the person in the video are deflected too large(lost part of main features of face), so that the detection module cannot find the face or the preprocessing module labels the feature points failed. Nevertheless, many continuous frames of one face can be obtained for the video, so the system eliminates some images which are not conducive to recognition is beneficial for false recognition rate reduction.

Conclusion

In this paper, a standardized method, using in practical face recognition system, is proposed. The method achieves face standardization to avoid mis-alignment by five steps, which are land mark localization, geometric rectification, core region extraction, size normalization and illumination normalization. Because of the controllable calculation amount, it can be realized in small practical equipment easily. In the experiment, a miniaturized face recognition system is built. The test results show that the recognition rate of this standardized method using the classical face recognition algorithms in practical face recognition environment has improved obviously. In future, we will extend the scope of angle change adaptation and achieve the standardization process in more terrible lighting conditions.

Acknowledgement

In this paper, the research was sponsored by the Fundamental Research Funds for the Central Universities (Project No. DUT16QY32) and Natural Foundation of Liaoning Province(Project No. 201601523).

References

- [1] Li Z, Xue L, Tan F. Face detection in complex background based on skin color features and improved AdaBoost algorithms[C]// Progress in Informatics and Computing (PIC), 2010 IEEE International Conference on. IEEE, 2010:723-727.
- [2] Tao H, Rui L, Mei-Juan Z. Face Recognition under Complex Conditions[C]// Electrical and Control Engineering (ICECE), 2010 International Conference on. IEEE, 2010:960-963.
- [3] Yi D, Lei Z, Li S Z. Towards Pose Robust Face Recognition[C]// Proceedings / CVPR, IEEE Computer Society Conference on Computer Vision and Pattern Recognition. IEEE Computer Society Conference on Computer Vision and Pattern Recognition. 2013:3539-3545.
- [4] Huang S M, Yang J F. Improved Principal Component Regression for Face Recognition Under Illumination Variations[J]. IEEE Signal Processing Letters, 2012, 19(4):179-182.
- [5] Hsieh C K, Lai S H, Chen Y C. An optical flow-based approach to robust face recognition under expression variations.[J]. IEEE Transactions on Image Processing A Publication of the IEEE Signal Processing Society, 2010, 19(1):233-40.
- [6] Han H, Shan S, Chen X, et al. A comparative study on illumination preprocessing in face recognition[J]. Pattern Recognition, 2013, 46(6):1691-1699.

- [7] Zhu Z, Luo P, Wang X, et al. Deep Learning Identity-Preserving Face Space[J]. 2013:113-120.
- [8] Lee K C, Ho J, Yang M H, et al. Video-based face recognition using probabilistic appearance manifolds[C]// IEEE Computer Society Conference on Computer Vision and Pattern Recognition. IEEE Computer Society, 2015:I-313-I-320 vol.1.
- [9] Jourabloo A, Liu X. Face Alignment Assisted by Head Pose Estimation[J]. Computer Science, 2015, 7.
- [10] Lai H, Xiao S, Cui Z, et al. Deep Cascaded Regression for Face Alignment[J]. Computer Science, 2015.
- [11] Wong Y, Chen S, Mau S, et al. Patch-based probabilistic image quality assessment for face selection and improved video-based face recognition [J]. Computer Science, 2014, 7255:74-81.
- [12] Cootes T F, Taylor C J, Cooper D H, et al. Active Shape Models-Their Training and Application[J]. Computer Vision & Image Understanding, 1995, 61(1):38–59.
- [13] Milborrow S, Nicolls F. Active shape models with SIFT descriptors and MARS[C]// International Conference on Computer Vision Theory and Applications . IEEE, 2014:380-387.
- [14] S. Milborrow. Building Stasm 4 Models. <http://www.milbo.users.sonic.net/stasm>, 2014.
- [15] Shan S, Gao W, Cao B, et al. Illumination Normalization for Robust Face Recognition Against Varying Lighting Conditions[C]// IEEE International Workshop on Analysis and Modeling of Faces and Gestures. IEEE Computer Society, 2003:157-157.
- [16] Lienhart R, Kuranov A, Pisarevsky V. Empirical Analysis of Detection Cascades of Boosted Classifiers for Rapid Object Detection[M]// Pattern Recognition. Springer Berlin Heidelberg, 2003:297-304.
- [17] Turk M A, Pentland A P. Face recognition using eigenfaces[C]// Cvpr, IEEE Computer Society Conference on Computer Vision and Pattern Recognition. IEEE Computer Society Conference on Computer Vision and Pattern Recognition. 2011:586-591.
- [18] Belhumeur P N, Hespanha J P, Kriegman D J. Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 1997, 19(7):711-720.