DOI: 10.2352/ISSN.2470-1173.2016.11.IMAWM-457

Face pose normalization and simulation methods based on multi-view face alignment

Changsong Liu[1,2,3] and Liting Wang[1]; 1Department of Electronic Engineering, Tsinghua University, Beijing 100084, China, 2State Key Laboratory of Intelligent Technology and Systems,3Tsinghua National Laboratory for Information Science and Technology (China)

Abstract

In this paper, we propose a face pose normalization and simulation methods based on multi-view face alignment that can enhance the performance of the face recognition algorithm towards large pose variation. The proposed method includes two steps: 1) multi-view face alignment, 2) face pose normalization and simulation methods. Multi-view face alignment algorithm is inspired by the design idea of the Supervised Descent Method (SDM) which is considered the state-of-the-art in face alignment. The proposed method modified the algorithm to adapt multi-view problems by changing the histogram of gradient feature to projection of gradient feature in order to adapt large pose variance. In addition, the feature scale also can be adaptive adjusted towards different part of face, for example, eyes, mouth, eyebrows, etc. Based on the multi-view face alignment results, 2D face normalization and simulation methods are proposed. Experimental results over many images with obvious pose changes have shown our method can significantly normalize the multi-view pose face and improve the accuracy of the existing common face recognition method when faces of probe sets have large pose variation.

Key words: face pose normalization; face simulation; multi-view face alignment; face recognition

Introduction

Over the past few decades, face recognition has become a popular area of research in pattern recognition and computer vision due to its wide range of commercial and law enforcement applications, such as biometric authentication, video surveillance, and information security. Until now, a great number of face recognition methods have been proposed and the system has been claimed widely used. However, pose variation is a challenge and the performance of current methods is still unsatisfactory. There is a wide variety of studies related to pose-robust face recognition when each identifier has only one frontal training sample [1]. Methods deal with large pose problems can be divided into three categories: the first category is the pose invariant feature extraction method. But the main disadvantage of this method is that it is difficult to extract the pose-invariant features. The second category is based on the multi-view face image solution. For example, the method in [2] extended the traditional subspace to multi-subspace. The main drawback of this method is it is very difficult to absolutely divide different face into accurate pose sets, and the errors of the pose estimation will reduce the performance of face recognition. The third category is the method based on 3D face model [3], based on 3D face model, the virtual face images are generated in solving the problem which has achieved good results. But a disadvantage is that the computational cost is large.

Robust face alignment is a pre-processing which could output a reliable and correct correspondence to face recognition normalization. Usually, we locate face landmarks and normalize face to extract features and train classifiers. Face alignment method under the most popular Active Shape Models (ASMs)[4] and the Active Appearance Models (AAMs) [5,6] has some disadvantages, the Supervised Descent Method (SDM)[7] refined the fixed shape PCA model, which is considered the state-of-the-art in face alignment. When it is utilized to multi-view and some large scale and more flexible situations, still, some improvement should be made to get a robust and reliable performance towards multi-view face pose changes. Face normalization relied on face alignment result, which is very important step for face recognition classifier training. There are kinds of face normalization towards pose changes by both 2D and 3D methods. How to normalize multi-view face image and get a better face recognition result is still a matter of concern.

The main objective of this proposed method is to design a face normalization and simulation method to improve the face recognition towards multi-view faces. The first proposed step is robust multi-view face alignment method. Though the Supervised Descent Method (SDM) provides a framework of face alignment, it still has some difficulties when it is utilized to multi-view and some large scale and more flexible situations. Compared with the generative feature, the discriminative feature shows good results in the past ASM and AAM framework. Under the SDM framework, the feature and feature scale could be chosen and discussed to improve the face alignment accuracy in multi-view cases. The second proposed step is face normalization method. Relied on face alignment results, multi-view face is normalized to frontal face template.

Proposed face pose normalization and simulation methods framework

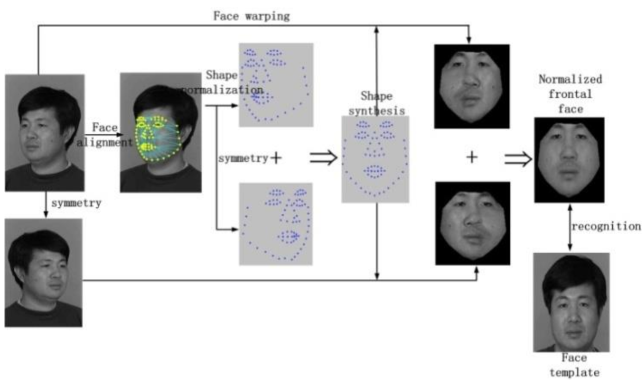

Fig.1. Proposed algorithm framework dealing with multi-view face pose variation

The basic idea proposed in this study is to deal with multi-view face recognition with the help of face alignment results. In face alignment step, each identifier's multi-view face is located by 88 feature points. Then, the shape and texture of both the multi-

©2016 Society for Imaging Science and Technology
DOI: 10.2352/ISSN.2470-1173.2016.11.IMAWM-457

view face and its symmetry face are normalized. By projecting them to the 2D frontal image plane, shape and texture are normalized to frontal view, thus, the testing face image was normalized and the face template was compared to the normalized frontal face as illustrated in figure.1.

The proposed framework consists of four modules, including 2D face alignment, shape normalization, texture normalization and face recognition. In this study, we focus on the former three modules. In this framework, the first module is an automatic 2D face alignment method. This 2D face fitting can estimate 88 facial feature points. The second and third modules are shape normalization and texture normalization, which relied on the results of face alignment results. The fourth module relates to face recognition algorithm by using the pose virtual face images, which is not the main topic of this study. Already, there have been great many contributions to face recognition using pose virtual face images, e.g. subspace learning methods, multiview-based learning methods, linear discriminant analysis and deep learning methods[8]. Therefore, we do not present the face recognition algorithm itself in detail.

Method

Face alignment method based on SDM

The traditional method of face alignment is based on parameter PCA model. Active shape model (ASM) and Active Appearance model (AAM) are usually used. The model is based on the coordinates of the feature points, under the constraints of local search and feature extraction, estimated feature points are closer to the true feature points step by step. Because of the existence of the model constraints, the two methods have some disadvantages. But as classic methods, AAM and ASM do not need a lot of training data, this is because these two algorithms have convergence feature point iterative learning. Xiong et al. proposed supervised descent method (SDM) as the feature points location algorithm in CVPR 2013. This algorithm is simple and effective. In quite a number of applications, SDM is gradually replacing the ASM and AAM algorithm and becomes the main stream. SDM is not only a feature points alignment algorithm, it can also solve the nonlinear optimization, especially in complex nonlinear optimization problems, the general Newton method needs to consume a large amount of computational resources, however, SDM combined with machine learning methods ensures the accuracy of the alignment at the same time, and greatly reduces the running time.

However, the experimental results show that the SDM algorithm is sensitive to initial value. Iterative expectation converges to the true feature points. For different faces, however, the path of convergence varies, and the paths of the extracted features are naturally different. When the amount of data is not enough, SDM algorithm could not reach higher accuracy.

The proposed method modified the face alignment algorithm to adapt multi-view problems by changing the histogram of gradient feature to projection of gradient feature in order to adapt large pose variance. As shown in figure 2, the projection of gradient feature has 8 axis (0,pi/4,pi/2, 3*pi/4, pi, 5*pi/4, 3*pi/2, 7*pi/4). Each gradient vector could be projected to nearest 2 axis. The feature is calculated by summing the intensity of each axis. Different with the histogram of gradient feature, the gradient vector is not calculated into different bins, instead, it is projected into the nearest two axis.

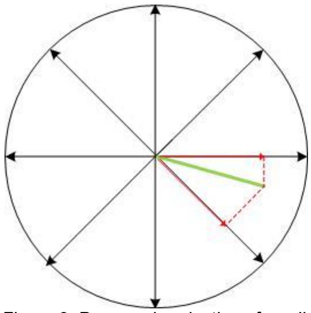

Figure 2. Proposed projection of gradient feature

The advantage of this feature is that it can accurately reconstruct the original gradient vector. This feature carries more information than the histogram of gradient feature. It shows to be a more accurate feature when it used to multi-view face alignment, as well as not losing the efficiency. Figure 3 shows the comparision of histogram of gradient feature and the proposed projection of gradient feature. Each pixel is extracted of local feature in the neighborhood of the gradient direction. It shows that the proposed projection of gradient feature's ability is better in description of details.

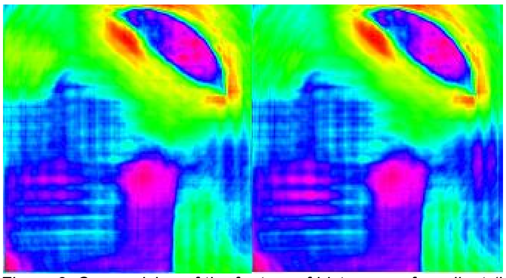

Figure 3. Comparision of the feature of histogram of gradient (left) and proposed projection of gradient (right)

In addition, the feature scale also can be adaptive adjusted towards different part of face, for example, eyes, mouth, etc. As Figure 4 illustrates different stage of optimization should use different scale of feature to train the transform matrix. In this paper, we calculate the feature scale adaptively to different label points to get better results.

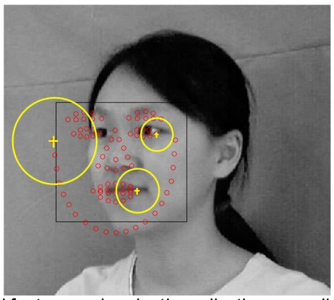

Figure 4. Proposed feature scale adaptive adjustion according to different part of face labels in order to adapt large pose variance.

Face pose normalization and simulation methods

Based on the multi-view face alignment results, 2D face normalization and simulation methods are proposed. Only one profile face image is used to get a simulated frontal face by using face symmetry information. As figure 1 illustrates, face symmetry

©2016 Society for Imaging Science and Technology
DOI: 10.2352/ISSN.2470-1173.2016.11.IMAWM-457

is utilized to add the information of the profile face image. The shape warping algorithm could be on 2D, and the texture will be simulated to reconstruct a frontal face.

Basic idea of normalization method

Basic idea of face normalization method is image warping by shape triangle transform. In figure 5, facial shape in each image is captured by recording the xy-coodinates of 34 facial landmarks. The images are then co-registered by morphing them to a fixed 34-point template using bi-cubic imterpolation.

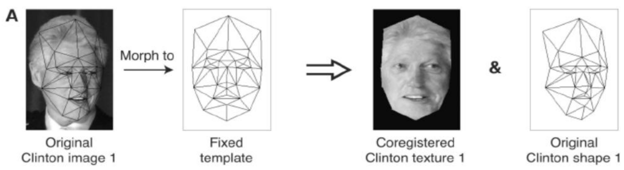

Figure 5: Basic idea of face normalization method[9]

In this paper, feature points number is 88, and the template is not a fixed template but calculated by the original shape and the symmetry shape. As shown in figure 6. The texture is simulated by warping result of the profile face and its symmetry face.

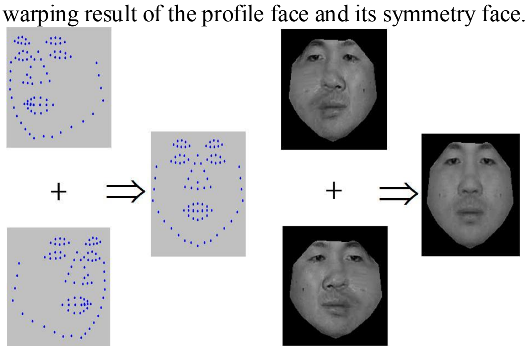

Figure 6: shape normalization and texture simulation

Shape normalization

As illustrated in figure 1, the multi-view face image to be normalized is I_o , its symmetry face image is I_f , I_f could be obtained by formula (1). The result of face alignment on I_o is the shape S_o , also, the shape S_f of the symmetry face I_f could be obtained by formula (2), here, W is the width of image, H is the height of image. L is the number of face shape's feature points. (l=88 in our work) .

$$I_f(x,y)=I_o(W-x,y);\ 1\le x\le W, 1\le y\le H \tag{1}$$

$$S_o=(x_1,y_1,x_2,y_2\cdots x_l,y_l)\ ;$$
$$S_f=(W-x_1,y_1,W-x_2,y_2\cdots W-x_l,y_l); \tag{2}$$

The normalized face shape is S_m , it could be calculated by formula (3), it is the mean shape of the original face shape and its symmetry face shape.

$$S_m=(S_o+S_f)/2 \tag{3}$$

Texture normalization

As illustrated in Fig.1, the multi-view face image to be normalized is I_o , its texture is T_o , its symmetry face texture is T_f , the normalized texture T_m could be calculated by:

$$T_m(\vec{x})=(T_o(W_{om}(\vec{x}))+T_f(W_{fm}(\vec{x})))/2 \tag{4}$$

W(•) is the piecewise affine transformation warping [5], as illustrated in figure 7, its transformation formula is (5) :

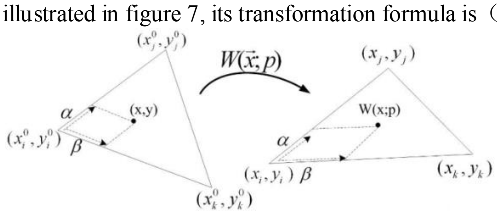

Figure 7: piecewise affine transformation warping [5]

$$\vec{x}=(x,y)=(x_i^0,y_i^0)+\alpha[(x_j^0,y_j^0)-(x_i^0,y_i^0)]$$
$$+\beta[(x_k^0,y_k^0)-(x_i^0,y_i^0)]$$
$$\alpha=\frac{(x-x_i^0)(y_k^0-y_i^0)-(y-y_i^0)(x_k^0-x_i^0)}{(x_j^0-x_i^0)(y_k^0-y_i^0)-(y_j^0-y_i^0)(x_k^0-x_i^0)}$$
$$\beta=\frac{(y-y_i^0)(x_j^0-x_i^0)-(x-x_i^0)(y_j^0-y_k^0)}{(x_j^0-x_i^0)(y_k^0-y_i^0)-(y_j^0-y_i^0)(x_k^0-x_i^0)}$$
$$W(\vec{x},p)=(x_i,y_i)+\alpha[(x_j,y_j)-(x_i,y_i)]$$
$$+\beta[(x_k,y_k)-(x_i,y_i)] \tag{5}$$

W_{om} represents simulation face from the original face, W_{fm} represents the face normalization from the symmetry face to the average face.

Experimental results

We experiment the proposed method on the publicly available FERET dataset[10]. Feret database includes faces viewed from 9 different poses as illustrated in figure 9. bi, bh, bg, bf, ba, be, bd, bc, bb. ba is the frontal face template.

We do face alignment to the multi-view face image and then simulate a frontal face for each pose. The example of simulated frontal face images are shown in the figure 8. It seems that normalization of multi-view faces and simulated frontal face show a very good quality.

Then, the proposed multi-view face normalization and simulation algorithm are applied to face recognition, compared with the recognition without any normalization to verify the proposed algorithm is helpful for large pose face recognition. Face recognition algorithm used here is the traditional Gabor features, linear discriminant analysis and the minimal distance classifiers.

DOI: 10.2352/ISSN.2470-1173.2016.11.IMAWM-457

Figure 8：Example of the results of face simulation

Figure 9 shows the results of face recognition rates toward multi-view face poses. In the multi-view situation, the results with face pose normalization receive better recognition rate than without pose normalization in the large pose sets. In the small pose sets (bf and be), however, the recognition rate has a little drop after the normalization. The reason is the simulation of texture may cause some noise and distortion.

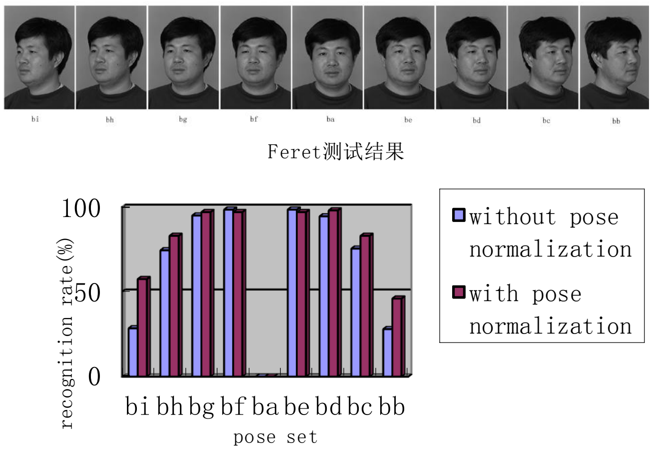

Figure 9 Results of face recognition

Conclusions

To deal with large pose changes in face recognition problems, the face normalization should be a matter of concern. This paper proposes a method of pre-processing of face normalization. Two important steps are proposed: One is the multi-view face alignment method under the framework of SDM and the other is the 2D face pose normalization and simulation method.

Differs from our previous published paper, which is under the framework of the most popular deformable models, the Active Shape Models (ASMs) and the Active Appearance Models (AAMs), proposed multi-view face alignment algorithm is inspired by the design idea of the Supervised Descent Method (SDM) which is considered the state-of-the-art in face alignment. We modified the fixed shape model and independent optimization to flexible global label joint optimization and we change the histogram of gradient feature to projection of gradient feature to get a better performance. In addition, the feature scale also can be adaptive adjusted according to different part of face labels in order to adapt large pose variance. Based on the face alignment results, 2D face normalization and simulation methods are proposed. Only one profile face image is used to get a simulation of frontal face by using face symmetry information which is different from our previous 3D pose simulation method [11].

Acknowledgment

This work was supported by the National Basic Research Program of China (973 program) under Grant No. 2013CB329403 and the National Natural Science Foundation of China under Grant No. 61471214.

References

[1] S. Du, R. Ward, "Face recognition under pose variations", Journal of the Franklin Institute, September, vol 343(6)(2006), pp.596-613

[2] A. Pentland, B. Moghaddam, and T. Starner, "View-based and modular eigenspaces for face recognition," in Proc. IEEE Conf. CVPR, 1994,pp. 84–91.

[3] V. Blanz and T. Vetter, "A morphable model for the synthesis of 3D faces," in Proc. SIGGRAPH, 1999, pp. 187–194.

[4] T. F. Cootes, C. Taylor, D. Cooper, and J. Graham. "Active shape models – their training and their applications". Computer Vision and Image Understanding, 61(1):38-59, January 1995

[5] I. Matthews and S. Baker, "Active appearance models revisited," International Journal of Computer Vision, Vol. 60, No. 2, November, 2004, pp. 135 - 164

[6] T. Cootes, G. Edwards, C. Taylor, "Active appearance models", IEEE Transactions on Pattern Analysis and Machine Intelligence 23 (6) ,2001, 681–685

[7] X. Xiong，F.D.l Torre."Supervised Descent Method and Its applications to Face Alignment", IEEE Conference on Computer Vision & Pattern Recognition, 2013, 9(4):532-539

[8] Y. Sun，X. Wang and X. Tang. "Deep Learning Face Representation from Predicting 10,000 Classes". IEEE Conference on Computer Vision & Pattern Recognition, 2014:1891-1898

[9] R. Jenkins and A. M. Burton, "100% Accuracy in Automatic Face Recognition", Science 25 January 2008, vol. 319. no. 5862, pp. 435

[10] P. J. Phillips, H. Moon, P. J. Rauss, and S. Rizvi, "The FERET evaluation methodology for face recognition algorithms", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 22, No. 10, October 2000.

[11] L.Wang, L.Ding, X.Ding and C.Fang, "2D face fitting-assisted 3D face reconstruction for pose-robust face recognition". Soft Comput. 15(3): 417-428 .2011.

Author Biography

Changsong Liu is an associate professor in the Department of Electronics engineering at Tsinghua University. He received the BS degree in both mechanics engineering and electronics engineering in 1992, Master degree in electronics engineering in 1995, Ph.D degree in 2007 from Tsinghua university, China. His fields of interests includes image processing, pattern recognition, nature language processing.

Liting Wang received her BS in electronic engineering from Tianjin University of China (2004) and her PhD in from Tsinghua University of China (2010). Her work has focused on the medical image analysis and face recognition algorithm.