3D Face Recognition by Collaborative Representation Based on Face Feature

Huaijuan Zang, Shu Zhan*, Mingjun Zhang, Jingjing Zhao, and Zhicheng Liang

School of Computer& Information, Hefei University of Technology, Hefei 230009, China
shu_zhan@hfut.edu.cn

Abstract. To overcome the crucial problem of illumination, facial expression and pose variations in 2D face recognition, a novel algorithm is proposed by fusing global feature based on depth images and local facial feature based on Gabor filters. These two features are fused by residual combined with collaborative representation. Firstly, this approach extracts Gabor and Global feature from 3D depth images, then fuses two features via collaborative representation algorithm. The fused residuals serve as ultimate difference metric. Finally, the minimum fused residual corresponds to correct subject. Extensive experiments on CIS and Texas databases verify that the proposed algorithm is effective and robust.

Keywords: collaborative representation, Gabor feature, global feature, 3D face recognition.

1 Introduction

Face recognition has become an active research topic in the field of biometric recognition. Although 2D face recognition technology is gradually mature, it is susceptible to illumination, pose and facial expression. Therefore, more and more researchers turn to the study of 3D face recognition.

In recent years, sparse representation is a concern in the area of signal processing. Wright et al [1] used sparse representation for face recognition in 2009. They presented sparse representation-based classification (SRC), which showed better robustness. It also achieved good classification results under the occlusion. But some researchers [2-4] have started to question for the role of l_1 norm in image classification. In 2011, Zhang et al [4] presented collaborative representation-based classification (CRC) for face recognition. It solved sparse coefficient by l_2 norm, using similar faces as training dictionary collaboratively, and obtained good recognition effects.

Gabor filter is widely used in the analysis of texture features and image recognition [5] for its good resolution in both time and frequency. In face recognition [6-7] Gabor feature can inhibit illumination, pose and facial expression. The local feature is easy to describe human face, which plays an important role in face recognition.

* Corresponding author.

Z. Sun et al. (Eds.): CCBR 2014, LNCS 8833, pp. 182–190, 2014.

3D face depth images have good robustness for illumination, expression and pose variations [8]. In this paper, feature selection is fulfilled face recognition combining with collaborative representation. Firstly, this approach extracts global and Gabor feature from 3D face depth images, and then it uses knowledge of collaborative representation to solve sparse coefficient. The minimum fused residual acquired from reconstruction, as the ultimate difference metric, is used to classify.

2 Collaborative Representation Based Classification(CRC)

Zhang et al [4, 9] think that when sparse representation used for face recognition, the key factor is constructing dictionary by multiple samples collaboratively. The dictionary by training samples is often less complete. We can make all training samples together constitute the dictionary. In order to reduce computational complexity, it solves the sparse coefficient by regularized least squares method. But the sparsity is not as strong as l_1 norm. So the classification criterion is improved. It greatly reduces complexity of the algorithm but has no recognition rate impairments. Suppose there are K classes of subjects, and let $X = [X_1, X_2, \cdots X_K]$ as training samples, a test sample for y , then

$$\hat{\rho} = \arg\min_{\rho} \left\{ \|y - X\rho\|_2^2 + \lambda \|\rho\|_2^2 \right\} \tag{1}$$

λ is regularization parameter, After mathematical derivation, the formula can be transformed into:

$$\hat{\rho} = \left(X^T X + \lambda \cdot I \right)^{-1} X^T y \tag{2}$$

Let $P = \left(X^T X + \lambda \cdot I \right)^{-1} X^T$. For a test face image y , it can just simply project y onto P , via $\hat{\rho} = Py$. This makes collaborative representation very fast. Then we reconstruct the test image, calculating the residual with the test image $\left\| y - X_i \hat{\rho}_i \right\|_2$, where X_i and $\hat{\rho}_i$ respectively correspond to the test image matrix and coefficient vector associated with class i . According to the principle of minimum residuals, it puts $\left\| \hat{\rho}_i \right\|_2$ into solving residuals, which provides more discriminant information for classification.

3 Feature Extraction

Gabor feature not only extracts identification components of the low frequency, but also well preserves the integrity information of face. The function of two-dimensional Gabor filters are defined as [6]:

$$\varphi_{u,v}(z)=\frac{\|k_{u,v}\|^2}{\sigma^2}\exp(-\frac{\|k_{u,v}\|^2\|z\|^2}{2\sigma^2})[\exp(ik_{u,v}z)-\exp(-\frac{\sigma^2}{2})] \tag{3}$$

Where $z=(x,y)$ denotes the pixel value of (x,y). u and v denote orientation and scale respectively. Wave vector is defined as $k_{u,v}=k_v e^{i\phi_u}$ with $k_v=k_{\max}/f^v$ and $\phi_u=\pi u/8$, $k_{\max}=\pi/2$ is the maximum frequency, and f is the spacing factor between kernels in the frequency domain ($f=\sqrt{2}$). σ determines the ratio of the Gaussian window width to wavelength($\sigma=2\pi$).

It can convolute a face image $I(z)$ and the Gabor filter $\varphi_{u,v}(z)$ to get Gabor feature $G_{u,v}(z)$:

$$G_{u,v}(z)=I(z)*\varphi_{u,v}(z)=M_{u,v}(z)\cdot\exp(i\theta_{u,v}(z)) \tag{4}$$

Where, $M_{u,v}(z)$ denotes amplitude and $\theta_{u,v}(z)$ denotes phase information. In this paper, five scales $v=\{0,1,2,3,4\}$ and eight directions $u=\{0,1,2,3,4,5,6,7\}$ are taken to obtain different directions and scales of Gabor feature, denoting $\chi=[a_{0,0}{}^T,a_{0,1}{}^T,\ldots,a_{4,7}{}^T]^T$. The extracted Gabor feature is conducted as input of classifier. From (4) we can see that $G_{u,v}(z)$ is a complex number. In this paper amplitude $M_{u,v}(z)$ is conducted as input because it contains the variation of image local energy which can be used as a measure of local feature.

4 Collaborative Representation Based on Face Feature

3D face depth image not only contains two-dimensional texture but also includes spatial information which is an inherent property of the face. The existence of expression, light, and occlusion will affect the accuracy of feature extraction in global feature. Local feature can divide images into different facial areas and treat them differently. Gabor feature has certain robustness to light, pose and facial expression. Gabor feature also has good spatial locality and orientation selectivity, which can keep local feature of the original data. But Gabor feature is very sensitive to occlusion, so the effect of identification is not very good under the occlusion. Due to the feature that CRC is insensitive to occlusion, an algorithm is presented based on face feature by collaborative representation. It plays advantageous in global and local feature respectively and also makes up its own shortcomings. Whether in the case of pose, facial expression or occlusion, it can greatly improve the recognition performance.

Flow diagram of collaborative representation based on face feature shows in figure 1:

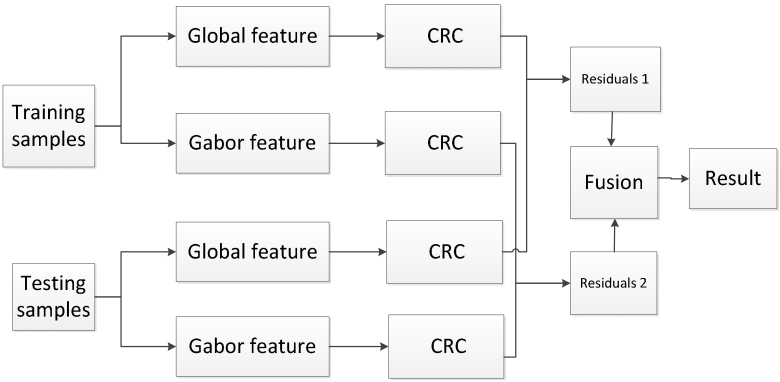

Fig. 1. Flow chart of 3D face recognition by collaborative representation based on face feature

Main steps of the method are as follows:

1) Training set is composed by 3D depth images. Assuming K classes samples, the training samples are $X = [X_1, X_2, \cdots, X_K]$, a test sample is y ;

2) Extract global feature from 3D depth images, then according to (3) and (4), extract Gabor feature which matrices are $G_{global} = [\mathrm{G}^{glo}{}_1, \mathrm{G}^{glo}{}_2, \cdots, \mathrm{G}^{glo}{}_k]$ and $G_{gabor} = [\mathrm{G}^{gab}{}_1, \mathrm{G}^{gab}{}_2, \cdots, \mathrm{G}^{gab}{}_k]$;

3）Project G_{gloK} and G_{gabK} into principal component analysis (PCA) subspace:

$$T^{glo}{}_r = W^T G^{glo},\quad T^{gab}{}_r = W^T G^{gab} \tag{5}$$

$$T^{glo}{}_t = W^T Y^{glo},\quad T^{gab}{}_t = W^T Y^{gab} \tag{6}$$

4) Training samples after reducing dimension constitute the training dictionary:

$$D_{glo} = [T^{glo}{}_{r1}, T^{glo}{}_{r2}, \cdots, T^{glo}{}_{rk}] \text{ and } D_{gab} = [T^{gab}{}_{r1}, T^{gab}{}_{r2}, \cdots, T^{gab}{}_{rk}];$$

5) Normalize each column of D_{glo} and D_{gab} , then project a test image to matrix P_1 and matrix P_2 respectively, so we can get the sparse coefficient vector $\alpha_1 = P_1 y$, $\alpha_2 = P_2 y$,where $P_1 = (D_{glo}{}^T D_{glo} + \lambda \cdot I)^{-1} D_{glo}{}^T$ and $P_2 = (D_{gab}{}^T D_{gab} + \lambda \cdot I)^{-1} D_{gab}{}^T$, I is unit matrix;

6) Calculate the residuals of various classes of samples:

$$e_i^{glo} = \frac{\left\| y - D^{glo}{}_i \alpha_1 \right\|_2}{\left\| \alpha_1 \right\|_2} \tag{7}$$

$$e_i^{gab} = \frac{\left\| y - D^{gab}{}_i \alpha_2 \right\|_2}{\left\| \alpha_2 \right\|_2} \tag{8}$$

7) Fused residual is the final measure:

$$e_i = \left| e^{glo}{}_i \right| + \left| e_i^{gab} \right| \tag{9}$$

8) $\mathrm{identity}(y) = \arg\min_i \{\mathrm{e}_i\}$

5 Experiments and Analysis

In order to verify effectiveness and robustness of the proposed method in dealing with illumination, facial expression and pose variations, extensive experiments were carried out on 2 databases: Texas [11] and real time imaging system face databases [10].

5.1 Real Time Imaging System Face Database

Real time imaging system face database is obtained by correlation image sensor. The database contains 77 pictures of 11 people with the size of 64*64. This database incorporates pose, facial expressions and occlusion(glasses / no glasses). Figure 2 shows some subjects from the database.

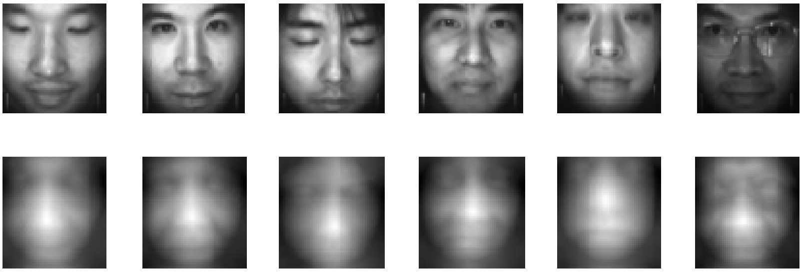

Fig. 2. Some human face intensity images and corresponding depth images in CIS

To some extent, the number of training samples affects recognition rates. 1,2,3,4,5,6 images are randomly selected as training samples of each individual respectively. The last ones are designated as probes. The selected images are 3D depth images. The experimental results are shown in Figure 3 at 4 dimensions.

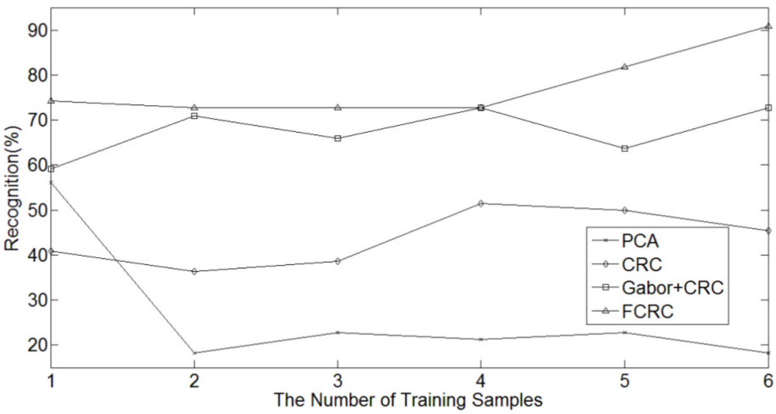

Fig. 3. Recognition rate curve with different number of training samples set in CIS

Since PCA is unable to overcome illumination, facial expression and pose, the recognition result is less effective. While CRC based on global and Gabor feature obtain a better recognition performance. Obviously, the number of training samples affects recognition rates. But it does not change the overall trend. Only the proposed algorithm with the increase numbers of training samples, the recognition rates kept increasing trend. They are significantly higher than several other algorithms. Experimental results show that the proposed method of the FCRC can get the optimal recognition results when choosing a different number of training samples.

In addition, for each subject, 3 depth images were selected as training samples, namely, the resting constitute probe samples. Figure 4 shows the recognition rate curve in different dimensions.

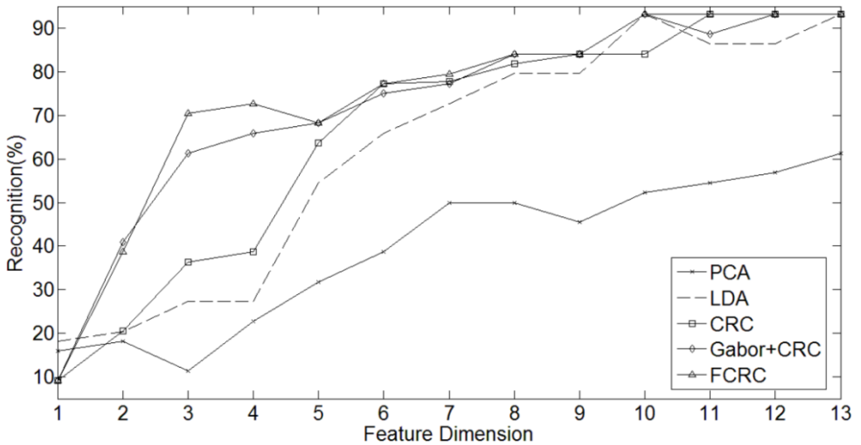

Fig. 4. Recognition rate curve with different feature dimensions in CIS

Linear discriminant analysis (LDA) firstly passes through PCA dimensionality reduction. From Figure 4, it can be observed that in most situations the proposed method outperforms other algorithms. It has reached 93.182% at dimension 10, and since then has maintained this level. While the other algorithms cover up to 93.182% at dimension 12 later, with exception PCA. Experimental results show that collaborative representation based on Gabor and global feature, and fused residuals serving as ultimate difference metric, the proposed method are superior to other algorithms, which proves the robustness of our algorithm.

5.2 Texas 3D Face Recognition Database

Texas 3D Face Recognition Database contains 1149 3D models of 118 adult human subjects. The facial expressions present are smiling or talking faces with open/closed mouths and/or closed eyes. The neutral faces are emotionless. Some human face images are showed in Figure 5.

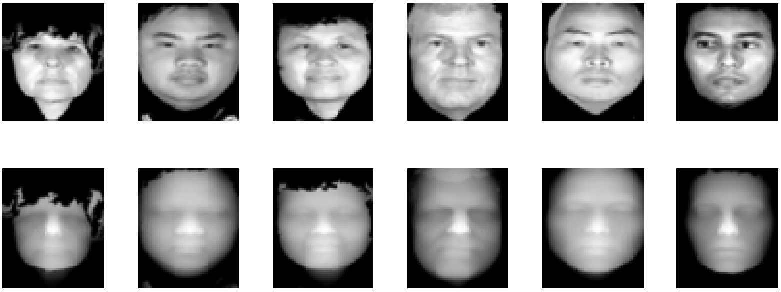

Fig. 5. Some human face intensity images and corresponding depth maps in Texas

1160 pictures of 116 classes are selected. Different training and test samples are set. The number of training samples of each class is 1,2,3,4,5,6,7,8,9, so the rest are used as probes. Figure 6 depicts results for different training samples at 16 dimensions.

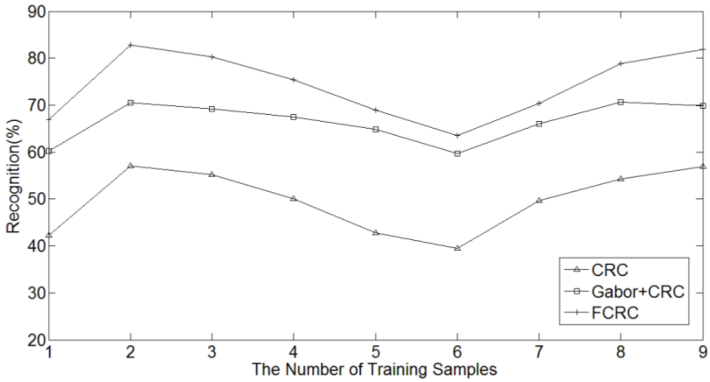

Fig. 6. Recognition rate curve with different size of training samples set in Texas

It can be seen from Figure 6, not training samples are more, and the recognition rates are higher. When the number increases, it will cause the redundancy of information. But no matter how many the training samples are, the proposed algorithm is more effective than other algorithms.

In order to demonstrate robustness of the algorithm in different dimensions, 1 image per person for training, a total of 116, and the other images are used as test samples, a total of 1044. Figure 7 shows the recognition rate curve in different dimensions.

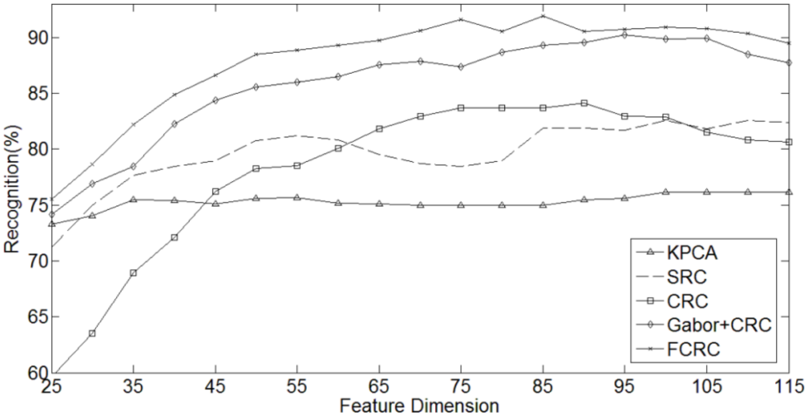

Fig. 7. Recognition rate curve with different feature dimensions in Texas

From Fig 7, with 66D between 112D features, FCRC achieves recognition rates above 90%, while Gabor-CRC reaches 90.134% at 94D features. After this dimension recognition, the rates drop below 90% almost as a whole. On the other hand, the best rates achieved by Kernel PCA (KPCA), SRC, original CRC and Gabor-CRC are 76.149%, 82.57%, 85.153% and 90.23%, FCRC outperforms others achieving a maximum recognition rate of 91.762%. It can be seen that compared with other algorithms, the proposed algorithm has higher recognition rates even when the number of training samples is small.

In order to fully evaluate the performance of the algorithm, the complexities of the various algorithms are analyzed, with Texas database as an example at 30 dimensions. The time consumed in the recognition phase is shown in table 1.

Table 1. Experimental comparison of our proposed method with other methods

Algorithm	TIME(s)
SRC	256
CRC	3
Gabor-CRC	4
FCRC	8

Since most of current face recognition system is offline, so we are comparing the time complexity of algorithms in the recognition phase. Since l_1 norm is more time-consuming to solve sparse coefficient, the time of SRC is the greatest. While CRC-based algorithms use regularized least square to solve the coefficient, time is far less than SRC. So the proposed method is effective.

6 Conclusions

In this paper, the algorithm by collaborative representation based on face feature is proposed. Firstly this approach extracts Gabor and Global feature from 3D depth images, then it fuses two features via collaborative representation algorithm. Since Gabor feature has good scale and orientation selectivity, CRC is insensitive to occlusion. Finally, the experimental results show that the proposed algorithm in different training samples and dimensions can effectively deal with occlusion, pose and expression variations.

Acknowledgments. This work is supported by National Natural Science Foundation of China(61371156).

References

1. Wright, J., Yang, A.Y., Ganesh, A., et al.: Robust face recognition via sparse representation. IEEE Transactions on Pattern Analysis and Machine Intelligence 31(2), 210–227 (2009)
2. Shi, Q., Eriksson, A., Van Den Hengel, A., et al.: Is face recognition really a compressive sensing problem? In: 2011 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 553–560. IEEE (2011)
3. Rigamonti, R., Brown, M.A., Lepetit, V.: Are sparse representations really relevant for image classification? In: 2011 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1545–1552. IEEE (2011)
4. Zhang, L., Yang, M., Feng, X.: Sparse representation or collaborative representation: Which helps face recognition? In: 2011 IEEE International Conference on Computer Vision (ICCV), pp. 471–478. IEEE (2011)
5. Hu, Z.P., Xu, B., Bai, Y.: Sparse representation for image recognition based on Gaborfeature set and discriminative dictionary learning. Journal of Image and Graphics 18(2), 189–194 (2013)
6. Yang, M., Zhang, L.: Gabor feature based sparse representation for face recognition with gabor occlusion dictionary. In: Daniilidis, K., Maragos, P., Paragios, N. (eds.) ECCV 2010, Part VI. LNCS, vol. 6316, pp. 448–461. Springer, Heidelberg (2010)
7. Shu, Z., Qixiang, Z., Jianguo, J., Ando, S.: 3D Face Recognition by Kernel Collaborative Representation Based on Gabor Feature. Acta Photonica Sinica 42(12), 1448–1453 (2013)
8. Drira, H., Ben Amor, B., Srivastava, A., Daoudi, M., Slama, R.: 3D face recognition under expressions, occlusions and pose variations. IEEE Transactions on Pattern Analysis and Machine Intelligence 35(9), 2270–2283 (2013)
9. Chi, Y., Porikli, F.: Classification and Boosting with Multiple Collaborative Representations. IEEE PAMI (2013)
10. Kurihara, T., Ono, N., Ando, S.: Surface orientation imager using three-phase amplitude-modulated illumination and correlation image sensor. In: Electronic Imaging 2003, pp. 95–102. International Society for Optics and Photonics (2003)
11. Gupta, S., Castleman, K.R., Markey, M.K., Bovik, A.C.: Texas 3D face recognition database. In: 2010 IEEE Southwest Symposium on Image Analysis& Interpretation (SSIAI), pp. 97–100. IEEE (2010)