Robust 3D Face Recognition in Uncontrolled Environments

Cheng Zhong, Zhenan Sun, Tieniu Tan and Zhaofeng He
National Laboratory of Pattern Recognition, Institute of Automation,
Chinese Academy of Sciences, Beijing, 100080, P.R.China
{czhong,znsun,tnt,zfhe}@nlpr.ia.ac.cn

Abstract

Most current 3D face recognition algorithms are designed based on the data collected in controlled situations, which leads to the un-guaranteed performance in practical systems. In this paper, we propose a Robust Local Log-Gabor Histograms (RLLGH) method to handle the uncontrolled problems encountered in 3D face recognition. In this challenging topic, large expressions and data noises are two main obstacles. To overcome the large expressions, we choose Log-Gabor features (LGF) to extract the distinctive and robust information embedded in 3D faces, which will be represented as 3D Log-Gabor faces. Data noises are summarized as distorted meshes, hair occlusions and misalignments. To overcome these problems, we introduce a Robust Local Histogram (RLH) strategy, which takes advantage of the robustness of the accurate local statistical information. The combination of LGF and RLH leads to RLLGH. The novelties of this paper come from 1) Our work aims at studying 3D face recognition performance in uncontrolled environments; 2) We find that embedding LGF into the LVC framework leads to robustness in handling large expression variations; 3) The RLH strategy gives a promising way to solve the data noises problem. Our experiments are based on the large expression subset in FRGC2.0 3D face database and the expression subset in CASIA 3D face database. Experimental results show the efficiency, robustness and generalization of our proposed method.

1. Introduction

Automatic identification of human faces is a very challenging research topic, which has gained much attention during the last few years. Most of this work, however, was focused on intensity or color images of faces [1]. Because of the robustness of 3D face recognition to pose and illumination variations, recently with the development of 3D acquisition system, a great deal of research effort has been devoted to this topic [2].

Earlier research in this field focused mainly on curvature analysis. Gordon et al. [3] adopted Gaussian and mean curvatures to characterize delicate features in 3D faces. Chua et al. [4] treated face recognition as a 3D non-rigid surface matching problem and divided the human face into rigid and non-rigid regions. They also proposed point-signature to describe 3D free-form surfaces. The rigid parts of faces were represented by point-signatures, and recognition was achieved by matching these signatures. Tanaka et al. [5] treated 3D face recognition as a 3D shape recognition problem of free-form curved surfaces. Each face in both input image and the model database was represented as an Extended Gaussian Image (EGI), and recognition was achieved by evaluating the similarities among these constructed EGIs. The methods based on curvature analysis make good use of 3D properties of the data, however, they usually require a high computation cost. In addition, because of the sensitivity of curvature based features, these methods also need high quality 3D face data.

Figure 1. The difficulties encountered in uncontrolled environments. The first pair shows the images in controlled situations, while the other pairs show the images with large expressions and data noises, which are the main obstacles in uncontrolled environments.

Recently because of the excellent performance of principal component analysis (PCA) in 3D face recognition [6], many appearance based methods have been adopted in this field. Lu et al. [7] constructed many 3D models as registered templates, then they matched 2.5D images (original 3D data) to these models using iterative closet point (ICP). Jamie et al. [8] adopted LGF to represent 3D faces, then the recognition was achieved by combining the scores in both spatial and frequency domain. Zhong et al. [9] intro-

978-1-4244-2243-2/08/$25.00 ©2008 IEEE

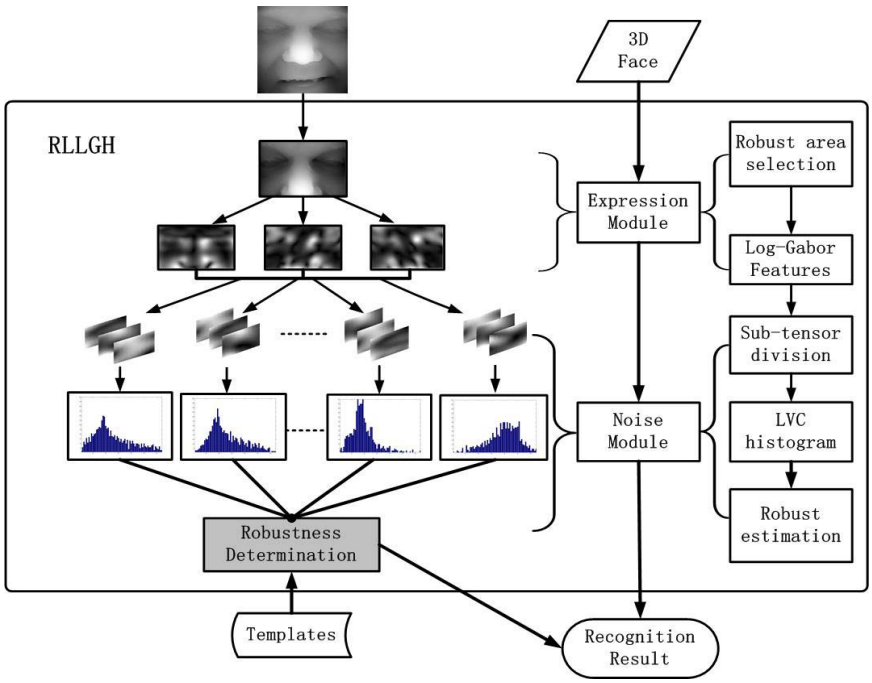

Figure 2. The flowchart of the proposed RLLGH method.

duced the Learned Visual Codebook (LVC) method into 3D face recognition. In their method, each 3D face was composed by some learned face codes, then according to this mapping, histogram vectors were obtained to represent the faces. Mian et al. [10] proposed a fusion system to handle the expression problem. In their system, three kinds of methods, spherical face representation, SIFT based matching and a modified ICP were combined to achieve the final recognition. Their results show the potential of appearance based methods to solve the expression problem in 3D face recognition. However, all the above methods [6] [7] [8] [9] [10] choose neutral faces as registered templates, which can be viewed as 3D face recognition in controlled environments. As Fig.1 shows, recognizing uncontrolled faces is much more difficult than recognizing neutral faces. Thus, can we accurately recognize the un-cooperative individuals when we only have their corresponding uncontrolled faces as registered templates?

In this paper, we introduce a RLLGH method aimed at handling this uncontrolled problem encountered in 3D face recognition. The flowchart of RLLGH is shown in Fig.2. Our method can be divided into two parts: to overcome the large expressions and to overcome the data noises. In expression module, first we select the relative robust face area to reduce the influence of the expressions, as shown in Fig.3. Then LGF is extracted from each face. The obtained 3D Log-Gabor faces are stored into a 3rd-order tensor. In noise module, first each tensor is divided into some local sub-tensors to reduce the complexity of noises. Then the LVC framework [9] is adopted to convert each sub-tensor into a LVC histogram vector. We also use robust estimation to choose the most robust local histograms for final recognition.

The contributions of this paper are as follows. First, a RLLGH method is proposed to recognize 3D faces in uncontrolled environments. Second, we embed the Log-Gabor filters into the LVC framework, which is efficient to characterize the individuals and robust to large expressions. Third, a RLH strategy is introduced to overcome the data noises. We also make a detailed comparison between RLLGH and some commonly used appearance based methods, such as PCA, Gabor features(GF), LGF, Local Binary Pattern (LBP), Local Gabor Binary Pattern Histogram Sequence (LGBPHS) and LVC [11] [12] [8] [13] [14] [9]. Experimental results show the effectiveness of our method to recognize 3D faces in uncontrolled environments.

The remainer of this paper is organized as follows. In Section 2, we give our motivation of 3D face recognition in uncontrolled environments. In Section 3, our RLLGH method is described in details. We describe our experimental results in Section 4. Finally, the paper is concluded in Section 5.

2. Motivation

FRGC2.0 database constructs an ideal platform for researchers to compare their algorithms. However, this database aims to test the influence of time lapse to recognition performance, which is the main difference of ROCI,

ROCII and ROCIII [15]. As to 3D face recognition, how to overcome the expression variations is also a challenging topic. Based on this, Geometrix gives a widely used protocol to divide the FRGC2.0 data into three subsets: neutral, small expression and large expression [16]. Neutral and small expression subsets can be viewed as 3D face recognition in controlled environments, while large expression subset can be viewed as 3D face recognition in un-cooperative environments. Recently, most work [10] is based on the neutral registered templates, such as neutral/neutral (neutral faces as gallery set and neutral faces as probe set), neutral/small expression (neutral faces as gallery set and small expression faces as probe set) and neutral/large expression (neutral faces as gallery set and large expression faces as probe set).

Because of the rapid development in 3D face recognition, the recognition performance of neutral faces has been improved substantially. We give the verification performances of RLLGH based on two protocols: ROCIII of FRGC2.0 and Geometrix neutral expression subset [15] [16]. In ROCIII of FRGC2.0, RLLGH can give a verification performance at 90.6% when FAR is 0.1%. In neutral subset of FRGC2.0, RLLGH can give a verification performance at 99.1% when FAR is 0.1%. These results are comparable to the state-of-art recognition performance. Therefore, in this paper we only aim at testing 3D face recognition performance in large expression subset [16].

3. Introduction of RLLGH

3D face recognition in uncontrolled environments has many unexpected factors. In this paper we address two issues: large expressions and data noises, as Eqn.1 shows.

$$I = P + E + N \tag{1}$$

where I is the face image, P refers to the characteristics of the person, E means the large expressions and N means the uncontrolled data noises. Our main target is to eliminate the expressions and noises embedded in the face image and preserve the main characteristics of the person. Therefore, our proposed RLLGH is composed of two modules, namely expression module and noise module, in which the former is to deal with the large expressions and the latter is to overcome the uncontrolled data noises. Next we will describe these two modules in details.

3.1. Expression module

This module contains two parts: robust area selection and Log-Gabor features extraction.

3.1.1 Robust area selection

All face areas will be influenced when large expressions happen. However, some facial components, such as nose, cheeks and eyebrows, have smaller variations than other facial components, such as mouth and chin. Therefore, to reduce the complexity of large expressions, we first select these relative robust areas as shown in Fig.3. To test the efficiency of this strategy, we define a rule called mean distance per pixel (MDP),

$$MDP = f(I_1 - I_2)/N \tag{2}$$

In Eqn.2, f is the distance function (L_1 is used here), I_1 and I_2 are the images to be matched, N is the number of pixels of I_1 and I_2. From Fig.2, we find that MDP is effectively reduced because of the robust area selection strategy.

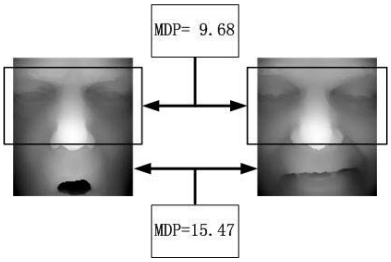

Figure 3. The relative robust face area. The MDP of selected face area (9.68) is smaller than that of the whole face area (15.47).

3.1.2 Log-Gabor features extraction

Fig.4(a) shows the Fourier spectrum in frequency domain of a 3D face. Although most of the energy is distributed in the low frequency area, they only reflect the global intensity distribution of the original image. And the more detailed discriminative information is embedded in the middle and high frequency area. Therefore, we need filters which can cover more bands in the middle and high frequency area.

Gabor filters have been used to many face recognition works because of its selectivity on orientation and spatial frequency [12] [9], as Eqn.3 shows.

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

where u and v denote the orientation and scale of Gabor kernels. However, as Fig.4(b) shows, the limited bandwidth of Gabor filters become the drawback of GF to represent 3D faces.

In [17] an alternative method was proposed to overcome this problem, which is the Log-Gabor filters, as Eqn.4, Eqn.5 and Eqn.6 show.

$$LGF(f_0, angle) = GTF_{LogGabor}(f_0) \odot Spread(angle) \tag{4}$$

$$GTF_{LogGabor}(f_0) = exp((-log(f/f_0)^2)/(2log(\sigma_1)^2)) \tag{5}$$

$$Spread(angle) = exp((-F(angle)^2)/(2\sigma_2^2)) \tag{6}$$

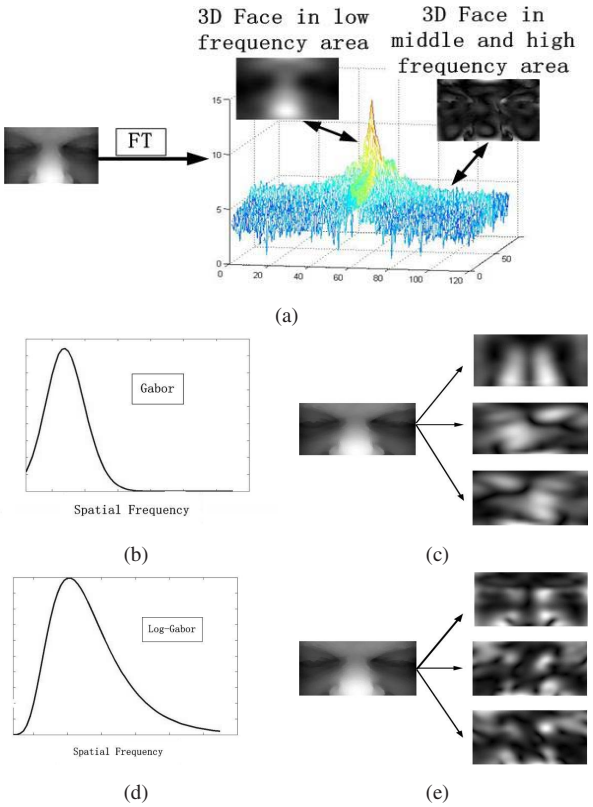

Figure 4. Gabor and Log-Gabor features of 3D face. (a): the Fourier spectrum of 3D face; (b)(c): the Gabor filter response and 3D Gabor Faces respectively; (d)(e): the Log-Gabor filter response and 3D Log-Gabor Faces respectively.

where f is the spatial frequency, f_0 is the center frequency of the filter, σ_1 and σ_2 are the standard deviation of Gaussian function, $angle$ is the polar angle in the frequency domain, F is a function which makes $Spread$ to be a Gaussian which is centered at the polar angle, $A \odot B$ means multiply each element in matrix A with its corresponding element in matrix B [18]. The main difference between Gabor and Log-Gabor is the Gaussian transfer function(GTF). When we convert Gabor representation (Eqn.3) into frequency domain, its GTF is shown as Eqn.7:

$$GTF_{Gabor}(f_0) = exp(-(f/f_0)^2)/(2(\sigma_1)^2)) \quad (7)$$

From the comparison of Eqn.5 and Eqn.7, we can find that in Log-Gabor filters, GTF is Gaussian when viewed on the logarithmic frequency scale. While in Gabor filters, GTF is Gaussian when viewed on the linear frequency scale. Therefore, the frequency response of Log-Gabor filters can cover more bands in the middle and high frequency area, which gives it the ability to capture more discriminative information embedded in the 3D faces, as shown Fig.4(d) and Fig.4(e).

The Log-Gabor representation of 3D face, called 3D Log-Gabor faces, as shown in Fig.4(e), can be obtained as follows:

$$3DLGF(f_0, angle) = IFT(FT(I) \odot LGF(f_0, angle)) \quad (8)$$

where I is a 3D face image, $LGF(f_0, angle)$ is the Log-Gabor filter with center frequency f_0 and polar angle $angle$, FT means the Fourier transform and IFT means the inversed Fourier transform, $A \odot B$ refers to multiply each element in matrix A with its corresponding element in matrix B. From the comparison of Fig.4(c) and Fig.4(e), it is clear that 3D Log-Gabor faces contain more detailed texture information than 3D Gabor faces. In our experiment we choose one center frequency (one scale) and three polar angles (three orientations) for the Log-Gabor filters. Therefore, each 3D face corresponds to three 3D Log-Gabor faces, which will be stored as a 3rd-order tensor.

3.2. Noise module

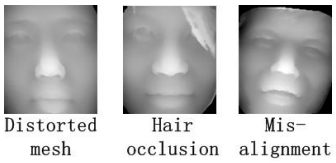

Figure 5. Three kinds of data noises: distorted mesh, hair occlusion and mis-alignment.

Besides large expression variations, data noises is another obstacle in uncontrolled environments. Here we summarize the data noises into three categories: Noises from 3D acquisition system, such as the distorted meshes; Noises from the subject, such as the hair occlusions; Noises from the registration, such as the mis-alignments. In our approach, RLH is adopted to overcome the data noises. The main steps of RLH can be divided into three parts: local sub-tensors division, LVC histograms construction and robust estimation.

3.2.1 Local sub-tensors division

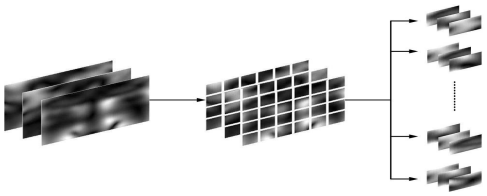

Figure 6. Dividing the 3D Log-Gabor faces into many local sub-tensors.

Local sub-tensors division is a kind of "divide and conquer" method. In this step, we divide the 3rd-order tensor, 3D Log-Gabor faces, into many local sub-tensors according to their spatial coordinates, as shown in Fig.6. It is the basic

step of our RLH strategy. For distorted meshes and mis-alignments, the dividing work largely reduces the complexity of noises. For hair occlusions, we assign the occluded face areas to only a few local sub-tensors, which will be filtered out in the next steps.

3.2.2 LVC histograms construction

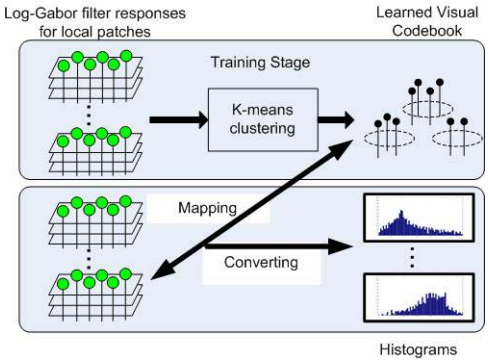

Figure 7. The LVC strategy.

As Fig.7 shows, we use k-means to learn the codes from each local sub-tensors in training, and these codes compose the learned visual codebook. Based on the constructed codebook, each local sub-tensor can be converted into a local histogram vector [9]. To handle distorted meshes and mis-alignments, the histogram representation will be more robust and efficient than appearance representation because of its statistical property.

3.2.3 Robust estimation

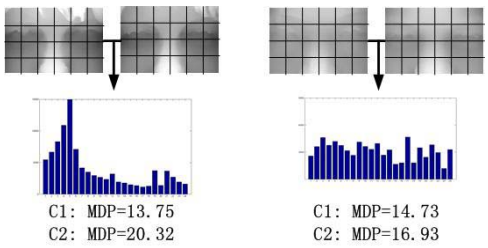

Figure 8. An example of robust estimation to overcome hair occlusion.

This step is to handle hair occlusions. In [19], several measures of robustness are mentioned in robust estimation, in which the most common is the breakdown points-the minimum fraction of outliers that can deteriorate the estimate accuracy. Local patches with hair occlusions are just such breakdown points when we compute the matching distance between individuals. An example of robust estimation is shown in Fig.8. Left pair is from the same individual, right pair is from different individuals. C1 is the case when we use the efficient local patches for comparison (when breakdown points are discarded), C2 is the case when we use all local patches for comparison. MDP is defined as Eqn.2. For C2, because of the influence of hair occlusions, the MDP of left pair is even larger than that of right pair, which is an error in verification. While for C1, thanks to the robust estimation (only the efficient local patches are adopted), the MDP of left pair is reduced from 20.32 to 13.75. Now it is smaller than that of right pair, which is only reduced from 16.93 to 14.73. Therefore, robust estimation can improve the recognition performance by overcoming hair occlusions.

To guarantee the robustness of our method, we need to combine the accurate distances between the local histograms to achieve recognition. Therefore, the local histograms with lower matching scores are viewed as breakdown points. Only the local histograms with higher matching scores are labeled as efficient and adopted for face recognition.

3.3. RLLGH framework

Our RLLGH framework can be concluded as follows:

- Expression module.
  - Select the relative robust face areas.
  - Extract Log-Gabor features to represent 3D faces, which called 3D Log-Gabor faces.
- Noise module.
  - Divide the 3D Log-Gabor faces into many local sub-tensors.
  - Apply LVC strategy to compute the statistical information of each sub-tensor and construct the LVC histograms corresponding to these local sub-tensors.
  - Select robust LVC histograms with higher matching scores for further recognition.

In our experiments, the selected robust face area is 80*120 pixels. After Log-Gabor transformation, the size of the obtained 3rd-order tensor is 80*120*3. In division, the size of each sub-tensor is 20*20*3. And 24(4*6*1) local sub-tensors are obtained in this step. In LVC, we adopt 64 clustering centers for each sub-tensor, so 24 64-dimensional histogram vectors are obtained to represent each 3D face. For robust estimation, we adopt L1 distance for matching and discard 4 histogram vectors with the lowest matching scores. Finally the matching scores of the remaining 20 efficient histogram vectors are combined to achieve recognition.

Although both LVC [9] and RLLGH adopt the learned visual codebook strategy, there are some notable differences between them.

(1) LVC [9] is constructed based on Gabor filter responses. While RLLGH is constructed based on Log-Gabor filter responses to overcome the expressions.

(2) In LVC [9], the clustering is based on the 20-dimensional vectors (5 scales and 4 orientations). While in RLLGH, the clustering is based on the 3-dimensional vectors (1 center frequency and 3 angles). Because clustering in higher dimensional space will introduce more noises into clustering centers, so the codes learned in RLLGH can more accurately reflect the original data distribution.

(3) In matching step, LVC [9] uses the all the local matching scores, while RLLGH only uses the scores from the local histograms which are estimated to be effective. Therefore, RLLGH is more robust to data noises.

3.4. Other methods used for comparison

PCA, GF, LGF, LBP, LGBPHS and LVC are commonly used appearance based methods in face recognition [11] [12] [8] [13] [14] [9]. Next we will make detailed comparisons between RLLGH and the above methods to show the efficiency, robustness and generalization of our proposed method for 3D face recognition.

4. Experimental results and discussion

Our proposed method is evaluated in terms of its robustness to handle the large expressions and data noises on two challenging databases: the large expression subset of FRGC2.0 3D Face Database [15] and the expression subset in CASIA 3D face database.

4.1. Experiments on large expression subset

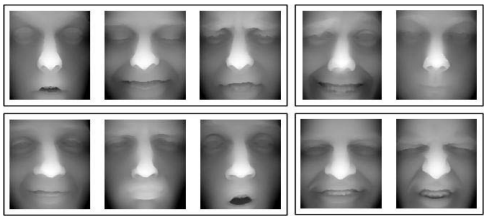

Figure 9. Some example images in large expression subset of FRGC2.0 3D Face Database. The images in the same black block are from the same person.

There are 742 faces from 311 individuals in the large expression subset of FRGC2.0 3D Face Database, which contain many kinds of expressions, such as laugh, big surprise, etc. [16]. To demonstrate the difficulty of this data set, we show some example images in Fig.9. Our experiments are arranged into two modes: verification mode and classification mode. In verification, we calculate a 742*742 similarity matrix. ROC curve and the verification rate when FAR = 1% are adopted to test the verification performance. In classification, we adopt the first image of each individual as the gallery set, which contains 311 faces. The remaining images are left as probe set, which contains 431 faces. CMS curve and the rank one score are adopted to test the classification performance. Our proposed method is compared with some widely used methods, such as PCA, GF, LGF, LBP, LGBPHS and LVC [11] [12] [8] [13] [14] [9]. Because some methods need the training stage, we use only 100 neutral faces from FRGC1.0 as the training set. The verification performance when FAR=1% is shown in Fig.10(a). The ROC curve is shown in Fig.10(b). The Rank1 classification rate is shown in Fig.10(c). The CMS curve is shown in Fig.10(d).

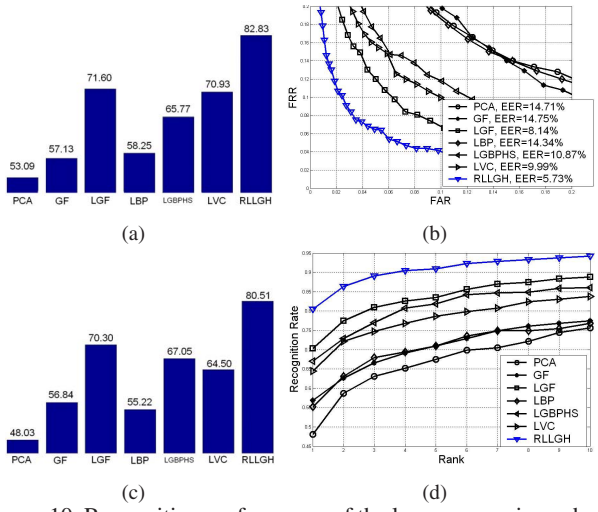

Figure 10. Recognition performance of the large expression subset from FRGC2.0 database in both verification mode and classification mode. (a) is the verification performance when FAR=1%. (b) is the ROC curve. (c) is the Rank1 classification performance. (d) is the CMS curve.

4.2. Experiments on expression subset

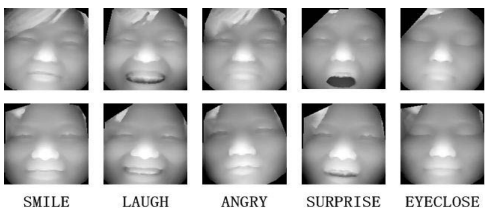

Figure 11. Some example images in expression subset of CASIA 3D Face Database. The images in the same row from the same person.

There are 615 faces from 123 individuals in the expression subset of CASIA 3D Face Database, which contain

many kinds of expressions, such as smile, laugh, surprise, etc. The expressions here is not that large as to the data in the first experiment, while there is more data noise in this data set. Some example images are shown in Fig.11. Our experiments are also arranged into two modes: verification mode and classification mode. In verification, we calculate a 615*615 similarity matrix. ROC curve and the verification rate when FAR = 1% are adopted to test the verification performance. In classification, we adopt the smile image of each individual as the gallery set, which contains 123 faces. The remaining images are left as probe set, which contains 492 faces. CMS curve and the rank one score are adopted to test the classification performance. Some widely used methods, such as PCA, GF, LGF, LBP, LGBPHS and LVC [11] [12] [8] [13] [14] [9] are adopted to compare with RLLGH. For the methods which need the training stage, we also use the same 100 neutral images from FRGC1.0 as the training set. The verification performance when FAR=1% is shown in Fig.12(a). The ROC curve is shown in Fig.12(b). The Rank1 classification rate is shown in Fig.12(c). The CMS curve is shown in Fig.12(d).

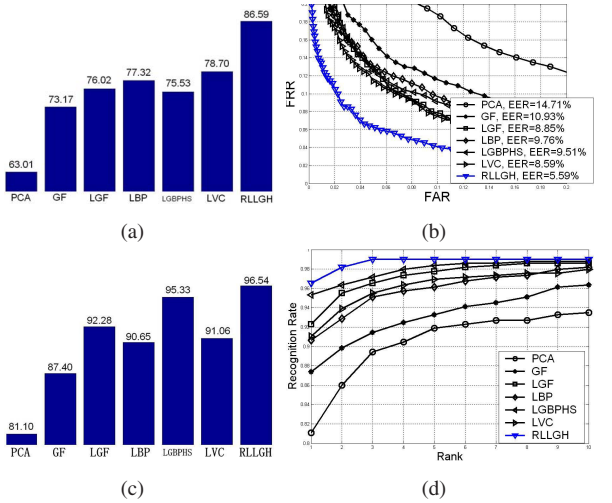

Figure 12. Recognition performance of the expression subset from CASIA database in both verification mode and classification mode. (a) is the verification performance when FAR=1%. (b) is the ROC curve. (c) is the Rank1 classification performance. (d) is the CMS curve.

4.3. Discussion

From the experimental results we can find that: first, LGF is more robust to expression variations than GF. Because the faces in the large expression subset of FRGC2.0 are manually aligned, thus the main obstacle in this experiment is large expression variations. The example images in Fig.9 demonstrate the difficulty of this data set. A notable result is that LGF obtains a substantially superior performance than PCA and GF, even LBP and LGBPHS, which shows its promising potential to handle the expression problem.

Second, our RLH strategy is a good solution to overcome the data noises. To test the robustness of algorithms, we manually rotate some faces to add the registration errors and select female faces with hair occlusions in the second experiment. Therefore, how to deal with the data noises become the main problem in this data set. For all the experimental results in both modes, RLLGH gives better recognition performance than LGF. Because $RLLGH = LGF + RLH$, thus the main difference between these two methods is the RLH strategy, which gives RLLGH the ability to handle the data noises problem.

Third, RLLGH provides an ideal framework to handle the problems encountered in uncontrolled environments for 3D face recognition. We summarize the difficulties in uncontrolled environments into two categories, large expressions and data noises. Then two methods, LGF for large expressions and RLH for data noises, are proposed to solve the above two problems respectively. LVC is adopted to combine LGF and RLH together, and this combination is RLLGH. Because it integrates the robustness and efficiency of LGF and RLH simultaneously, RLLGH obtains the best recognition results, both in verification mode and in classification mode, which shows its potential to be applied in uncontrolled environments.

Another outstanding advantage of RLLGH is its excellent generalization ability. In our experiments, we use only 100 neutral faces in FRGC1.0 as the training set. However, these experiments are carried out on two challenging expression data sets, which are shown in Fig.9 and Fig.11 respectively. There are no neutral images in the above two data sets. Therefore, our experiments can not only be viewed as cross-expressions (neutral images for training and expression images for testing), but also cross-databases (FRGC images for training and CASIA images for testing in the second experiment). Even worse, there are much more data noises in the testing data. In such difficult situations, RLLGH still gives the best recognition performance in all of the experiments. The strong learning ability makes it to be a good choice when we only have a limited number of training images on hand.

4.4. Future work

Although RLLGH shows its potential to solve the uncontrolled problem, its recognition performance can not be guaranteed when the expression is too large. The main reason is that RLLGH still belongs to the appearance based methods. Although we can choose the most robust features to represent the faces, such as LGF, we can not handle all the expressions. Some model based methods, such as deformable model [20] [21], can fit the learned model to the desired expressions. In such a way, their performance is

guaranteed whatever expression happens. In the future, we plan to solve the expression problems using the model fitting strategy.

5. Conclusion

Our method has been proposed to address a challenging topic, 3D face recognition in uncontrolled environments, in which large expressions and data noises are two main obstacles. We adopt LGF to overcome the large expressions and RLH to overcome the data noises. We also effectively integrate these two methods using the LVC framework, which leads to RLLGH. The promising recognition performance shows its potential to deal with the uncontrolled factors encountered in 3D face recognition.

Acknowledgement

This work is supported by the National Basic Research Program of China (Grant No. 2004CB318100), the National Natural Science Foundation of China (Grant No. 60736018, 60335010, 60702024, 60723005), the National Hi-Tech Research and Development Program of China (Grant No.2006AA01Z193, 2007AA01Z162), and the Chinese Academy of Sciences.

References

[1] W. Zhao, R. Chellappa, and A. Rosenfeld. Face recognition: a literature survey. *ACM Computing Surveys*, 35:399–458, 2003. 1

[2] K. W. Bowyer, K. I. Chang, and P. J. Flynn. A survey of approaches and challenges in 3d and multi-modal 3d+2d face recognition. *Computer Vision and Image Understanding*, 101:1–15, 2006. 1

[3] G. Gordon. Face recognition based on depth maps and surface curvature. *SPIE*, pages 234–274, 1991. 1

[4] C. Chua, F. Han, and Y. Ho. 3d human face recognition using point signature. *IEEE Inetrnational Conference on Automatic Face and Gesture Recognition*, pages 233–238, 2000. 1

[5] Hiromi T. Tanaka, Masaki Ikeda, and Hisako Chiaki. Curvature-based face surface recognition using spherical correlation principal directions for curved object recognition. *International Conference on Pattern Recognition*, pages 638 – 642, 1996. 1

[6] K. I. Chang, K. W. Bowyer, and P. J. Flynn. An evaluation of multi-modal 2d+3d face biometrics. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(4):619–624, 2005. 1, 2

[7] X. Lu, A. K. Jain, and D. Colbry. Matching 2.5d face scans to 3d models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(1):31–36, 2006. 1, 2

[8] Jamie Cook, Vinod Chandran, and Clinton Fookes. 3d face recognition using log-gabor templates. *BMVC*, 2006. 1, 2, 6, 7

[9] Cheng Zhong, Zhenan Sun, and Tieniu Tan. Robust 3d face recognition using learned visual codebook. *IEEE Conference on Computer Vision and Pattern Recognition*, 2007. 1, 2, 3, 5, 6, 7

[10] Ajmal S. Mian, Mohammed Bennamoun, and Robyn Owens. An efficient multimodal 2d-3d hybrid approach to automatic face recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(11):1927–1943, 2007. 2, 3

[11] M. Turk and A. Pentland. Eigenfaces for recognition. *J. Cognitive Neuroscience*, (1), 1991. 2, 6, 7

[12] Chengjun Liu and H. Wechsler. Gabor feature based classification using the enhanced fisher linear discriminant model for face recognition. *IEEE Transactions on Image Processing*, 11(4):467–476, 2002. 2, 3, 6, 7

[13] T. Ojala, M. Pietikainen, and T. Maenpaa. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(7):971–987, 2002. 2, 6, 7

[14] Wenchao Zhang, Shiguang Shan, Wen Gao, Xilin Chen, and Hongming Zhang. Local gabor binary pattern histogram sequence (lgbphs): A novel non-statistical model for face representation and recognition. *International Conference on Computer Vision*, pages 786 – 791, 2005. 2, 6, 7

[15] P. J. Phillips, P. J. Flynn, T. Scruggs, K. W. Bowyer, J. Chang, K. Hoffman, J. Marques, J. Min, and W. Worek. Overview of the face recognition grand challenge. *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 2005. 3, 6

[16] Thomas Maurer, David Guigonis, Igor Maslov, Bastien Pesenti, Alexei Tsaregorodtsev, David West, and Gerard Medioni. Performance of geometrix $activeid^{TM}$ 3d face recognition engine on the frgc data. *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 2005. 3, 6

[17] D. Fields. Relations between the statistics of natural images and the response properties of cortical cells. *Journal of Optical Society of America*, 4(12):2379–2394, 1987. 3

[18] http://www.csse.uwa.edu.au/ pk/research/matlabfns/PhaseCongruency/Docs/convexpl.html. 4

[19] Charles V. Stewart. Robust parameter estimation in computer vision. *SIAM Review*, 41(3):513 – 537, 1999. 5

[20] X. Lu and A. K. Jain. Deformable modeling for robust 3d face matching. *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 2006. 7

[21] Ioannis A. Kakadiaris, G. Passalis, G. Toderici, M. N. Murtuza, Y. Lu, N. Karampatziakis, and T. Theoharis. Three-dimensional face recognition in the presence of facial expressions: An annotated deformable model approach. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(4):640–649, 2007. 7