

Selection and Combination of Local Gabor Classifiers for Robust Face Verification

Nuri Murat Arar[†]
Bogazici University

nuri.arar@boun.edu.tr

Hua Gao
KIT*

hua.gao@kit.edu

Hazım Kemal Ekenel
KIT, ITU[‡]

ekenel@itu.edu.tr

Lale Akarun
Bogazici University

akarun@boun.edu.tr

Abstract

Gabor features have been extensively used for facial image analysis due to their powerful representation capabilities. This paper focuses on selecting and combining multiple Gabor classifiers that are trained on, for example, different scales and local regions. The system exploits curvature Gabor features in addition to conventional Gabor features. Final classifier is obtained by combining selected classifiers using Sequential Forward Floating Search-based selection mechanism. In addition, we combine classifiers trained on different local representations at score-level by learning the weights with partial least square regression. The system is evaluated on Face Recognition Grand Challenge (FRGC) version 2.0 Experiment 4. The proposed system achieves 94.16% verification rate @ 0.1% FAR, which is the highest accuracy reported on this experiment so far in the literature.

1. Introduction

Face recognition (FR) attracts attention from a wide variety of fields, from psychology and neuroscience to computer science. The problem has become popular in computer vision due to both its scientific challenges and the abundance of potential applications such as biometric identity authentication, video surveillance, smart cards, human computer interfaces and e-services including e-home, e-commerce and e-banking.

Within the past two decades, many FR algorithms have been proposed. A widely recognized traditional method is

subspace-based FR such as Eigenface (PCA) [20], Fisherface (FLD) [3] and Independent Component Analysis (ICA) [2]. These methods attempt to represent a face image as a linear combination of the basis images. Additionally, many researchers propose to solve the problem in the frequency domain by transforming a face image to the frequency domain and then, representing the image using the coefficients in the low frequency band. These spatial-frequency methods use Fourier transform [11], [9] and discrete cosine transform (DCT) [7], [17] to extract the features. On the other hand, Gabor wavelet based methods have been preferred due to their power of representation. Typical Gabor wavelet based methods include Elastic Bunch Graph Matching (EBGM) [21], and Gabor representation with Kernel Fisher Analysis (KFA) [10]. Later, researchers [18], [19], [12] tended to combine different types of features and classifiers with the help of increasing computational power.

Most of the studies using Gabor wavelets generate multiple features and classifiers based on different scales, orientations, and local regions -in case a local appearance-based approach is applied. Although feature selection problem in high-dimensional Gabor feature space has been addressed before, no studies have been conducted on selecting and combining multiple Gabor classifiers that are trained on different scales and local regions. Previous studies have performed fusion mainly by simple sum rule and without utilizing classifier selection. However, these different classifiers often contain complementary information that can improve the performance, if it is exploited efficiently. In this paper, we focus on this problem and employ classifier selection and fusion to develop a robust face verification system.

We propose a local appearance-based FR framework which uses Curvature Gabor (CG) features on spatially partitioned non-overlapping blocks on high-resolution (HR) face images, and then downscale the resulting Gabor images by averaging on grids. We use various parameter combinations in order to acquire complementary information from different curvature and scale spaces. Then, we apply principal component analysis (PCA) followed by linear discriminant analysis (LDA) independently on each block to ob-

* Karlsruhe Institute of Technology

‡ Istanbul Technical University

† Nuri Murat Arar is now with the Signal Processing Lab (LTS5) at the École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland.

This work has been supported by Tubitak-BMBF joint project 108E161, and OSEO, French State agency for innovation, as part of the Quaero Programme. It is also funded by the "Concept for the Future" of KIT within the framework of the German Excellence Initiative.

tain CG classifiers besides further dimensionality reduction. A CG classifier for an entire face image is the combination of block decisions with sum rule. After obtaining CG classifiers, we improve each classifier’s performance with log-likelihood ratio (LLR) based score conversion. Final classifier is obtained by combining selected classifiers using Sequential Forward Floating Search (SFFS)-based selection mechanism. In addition, we combine classifiers trained on different local representations at score-level by learning the weights with partial least square regression (PLSR). The proposed framework is evaluated on FRGC version 2.0 Exp. 4, which is a challenging experiment that requires matching of images collected under controlled and uncontrolled conditions. We achieved 94.16% verification rate (VR) @ 0.1% false acceptance rate (FAR), which is the highest performance reported on this experiment so far in the literature.

The rest of the paper is organized as follows: Related work is reviewed in Section 2, and the proposed method is introduced in Section 3. Experimental results and discussions are given in Section 4; followed by conclusion in Section 5.

2. Previous Work

The release of FRGC database version 2.0 has sped up the research in FR [15]. The most successful FR approaches combine different types of features in different scales that enable acquiring complementary information from each feature type, and scale. For example, Tan and Triggs [19] proposed to combine Gabor and LBP features. Su et. al. [18] suggested a hierarchical framework combining holistic Fourier classifier and local Gabor classifiers. Local classifiers are obtained from local Gabor patches on HR images while holistic classifier is obtained from Fourier features on low-resolution (LR) images. Liu and Liu [12] proposed a hybrid color-space based method. They combine local Gabor patch classifiers obtained from HR images, LBP and DCT classifiers obtained from LR images. Gao et al. [6] introduced multi-resolution local appearance based method. They combine local Gabor classifiers obtained from HR, LR and medium resolution face images.

In addition to all above methods, curvelet-based methods [13], [8] are presented to show that curvelets have better directional decomposition capabilities than wavelets for face images due to the characteristics of facial components. Recently, Hwang et. al. [8] introduced extended curvature Gabor wavelet (ECG)-based FR method on LR images. They extend conventional Gabor wavelet by including smaller frequencies, multiple Gaussian sizes, and different curvature parameters for object representations. They employ AdaBoost for feature selection in order to reduce the dimensionality of the extended features. Then, they learn ECG classifiers by applying LDA on each extended filter. Lastly, they merge a bunch of these classifiers using LLR-

based fusion.

3. Proposed Method

3.1. Gabor Wavelet

Gabor wavelets were inspired by 2D receptive field profiles of the mammalian cortical simple cells. They have been recognized as one of the most successful local feature extraction methods for face representation. They show desirable characteristics of spatial locality and orientation selectivity, and are optimally localized in the space and frequency domains.

The conventional Gabor wavelet is defined as follows:

$$\psi(\vec{x}; \nu, \mu) = \frac{k_{\nu, \mu}^2}{\sigma^2} e^{-\frac{k_{\nu, \mu}^2 \|\vec{x}\|^2}{2\sigma^2}} [e^{ik_{\nu, \mu} \vec{x}} - e^{-\frac{\sigma^2}{2}}] \quad (1)$$

where $\vec{x} = \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \cos \phi + y \sin \phi \\ -x \sin \phi + y \cos \phi \end{pmatrix}$ and $e^{ik_{\nu, \mu} \vec{x}}$ is the oscillatory wave function, whose real part and imaginary parts are the cosine and sine functions, respectively. ν controls the scale of Gabor wavelet, which mainly determines the center of the Gabor filter in the frequency domain; μ controls the orientation of the Gabor filters. In general, Gabor wavelet with the following parameters is used for face recognition: five scales $\nu \in \{0, 1, 2, 3, 4\}$ and eight orientations $\mu \in \{0, 1, \dots, 7\}$ with Gaussian window size $\sigma = 2\pi$.

3.2. Curvature Gabor (CG) Wavelets

A typical face image contains facial components such as eyes, nose, cheeks, lips, and eyebrows. These components show curved characteristics rather than straight ones. Therefore, it is natural to represent a face image with curvature kernels as well as straight ones.

Using CG wavelets, which are called as banana wavelets, is one way of modeling curve-like features of a face image. They are obtained by adding a curvature parameter to the conventional Gabor formulation [14]. The CG wavelets are defined as follows:

$$\psi(\vec{x}; \nu, \mu) = \frac{k_{\nu, \mu}^2}{\sigma^2} e^{-\frac{k_{\nu, \mu}^2 \|\vec{x}\|^2}{2\sigma^2}} [e^{ik_{\nu, \mu} \vec{x}} - e^{-\frac{\sigma^2}{2}}] \quad (2)$$

$$\vec{x} = \begin{pmatrix} \hat{x} \\ \hat{y} \end{pmatrix} = \begin{pmatrix} x \cos \phi + y \sin \phi + c(-x \sin \phi + y \cos \phi)^2 \\ -x \sin \phi + y \cos \phi \end{pmatrix} \quad (3)$$

where the curvature ratio $c \in \{0, 0.05, 0.1, 0.15, 0.2\}$.

CG wavelets do not have the orientation symmetry as in conventional Gabor wavelet as shown in Fig. 1. Therefore, if the curvature degree is not zero, the number of orientations used in CG wavelets increases to 16. For $c = 0$, it is simply the conventional Gabor wavelet.

In addition, CG wavelets use different Gaussian sizes, $\sigma \in \{0.5\pi, \pi, 1.5\pi, 2\pi\}$, for scale space utilization. Considering these parameter settings, we have 20 ($= 5 \times 4$)

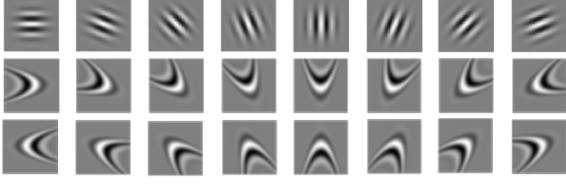


Figure 1. Orientation asymmetry in CG wavelets $c = 0.1$ (middle and bottom row) unlike conventional Gabor wavelet (top row).

different CG wavelets (Fig. 2) each consisting of 5×16 or 5×8 discrete filters, resulting in 1440 filters ($= 5 \times 16 \times 4 \times 4 + 5 \times 8 \times 4$).

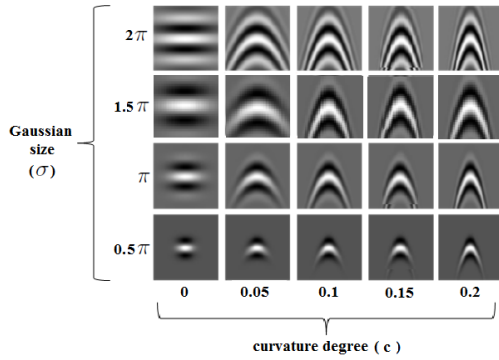


Figure 2. CG wavelets' real part representation with changing curvature degrees and Gaussian sizes.

To sum up, in addition to conventional Gabor wavelets which are good at representing coarse and straight structures, CG wavelets provide a way to model finer features with smaller Gaussian sizes e.g. $\sigma = \{0.5\pi, \pi, 1.5\pi\}$, and curvature features with different curvature degrees e.g. $c = \{0.05, 0.1, 1.5, 2\}$. CG wavelet modeling is supposed to have stronger representation power of modeling facial structures than the conventional Gabor wavelet.

3.3. Feature Extraction

Face registration is a crucial step for accurate FR. Registration parameters affect both the face region extracted and the proportions of facial components. These parameters can be adjusted to achieve better performance. We use eye centers and inter-ocular distance as registration parameters. We consider the trade-off between information content and noise while determining the alignment parameters. Face registration result on a sample input image is shown in Fig. 3.

A critical issue in Gabor feature extraction is the selection of the resolution. Many face characteristics are fine features; and would benefit from HR. However, using HR images brings two disadvantages: First, dimensionality increase brings computational load. Secondly, performance

of the system becomes very sensitive to landmark localization and face registration errors. To prevent these setbacks and still benefit from HR, we first perform full convolution of a CG wavelet with the HR face image to obtain the CG magnitude images in different scales and orientations. We then spatially partition each CG magnitude image into a number of non-overlapping local blocks of size $m \times m$ (Fig. 3), and downscale the CG features in each block, by effectively averaging the CG images. The spatial partitioning overcomes the problem of local information loss, makes the system robust to registration errors, and it also provides a relatively lower dimensionality.

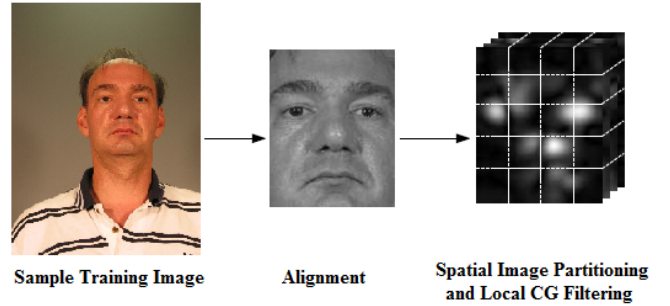


Figure 3. Feature extraction steps.

3.4. CG Classifiers

After obtaining CG features, we first perform Z-normalization, defined in equation (4), prior to the subspace analysis in order to centralize the data and normalize the variance.

$$z = \hat{x} = \frac{x - \mu}{\sigma} \quad (4)$$

where x is a feature vector, μ is the mean vector, and σ is the standard deviation vector.

After the normalization, we perform PCLDA, that is applying PCA followed by LDA on each block's normalized features independently. Applying PCLDA results in N local block classifiers based on nearest neighbor with normalized cross correlation as distance metric. Then, the decision of each block classifier, γ_i , is accumulated to form a single image classifier as shown in equation (5).

$$\Gamma = \sum_{i=1}^N \gamma_i \quad (5)$$

As stated in Section 3.2., there are 20 parameter configurations by using combinations of 5 different curvature degrees, $c \in \{0, 0.05, 0.1, 0.15, 0.2\}$ and 4 different Gaussian sizes, $\sigma \in \{0.5\pi, \pi, 1.5\pi, 2\pi\}$. For each parameter setup, process described above is applied to generate 20 CG classifiers.

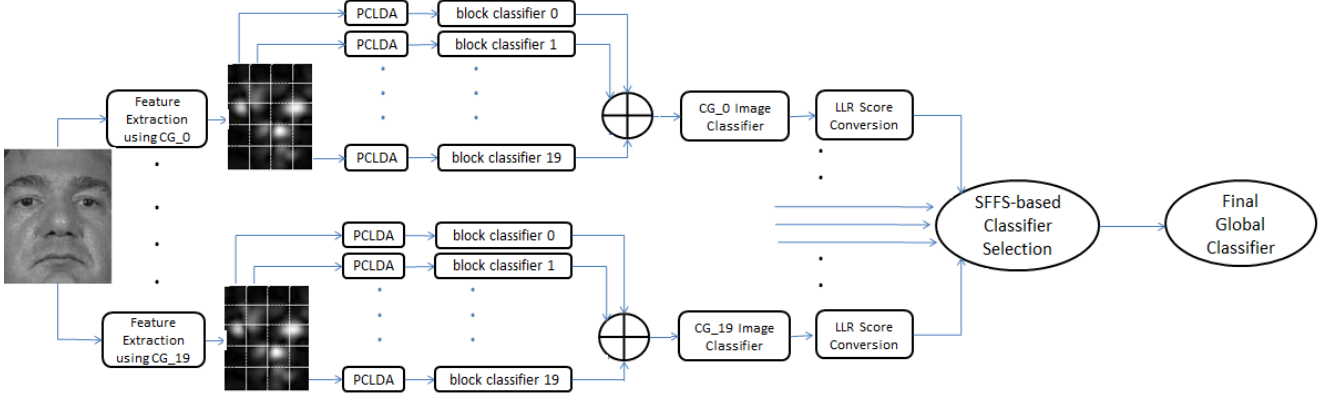


Figure 4. System block diagram.

3.5. Selection of CG Classifiers

Each of 20 CG wavelets is good at representing some particular features; therefore, some classifiers have complementary information when combined with others. This means that fusing some of the CG classifiers can increase classification performance. The optimal combination can be found by exhaustive search, but this requires 2^{20} trials which is practically impossible. Another method is to perform empirical selection that is simply fusing the classifiers with the highest VRs. However, it is important to note that the VR of a classifier alone is not the best criterion for determining the classifiers to be fused because the idea is to find the best combination, providing as much complementary information as possible. So we adapt a widely used feature selection algorithm, Sequential Forward Floating Search (SFFS) [16], for exploring this complementarity.

SFFS-based classifier selection algorithm starts from a null classifier set, and for each step, it adds the classifier that maximizes the VR on the training set when combined with the classifiers that have already been selected. The algorithm also verifies the possibility of VR improvement if a classifier is excluded.

SFFS-based classifier selection algorithm is as follows:

1. Initialization: $Y_0 = \{\emptyset\}$, $VR_{new} = 0$, $k = 0$

2. While $k < 20$ and $\Psi(Y_k) \leq VR_{new}$

- Inclusion of a classifier:

$$\Gamma_{LLR}^+ = \operatorname{argmax}_{\Gamma_{LLR} \notin Y_k} \Psi(Y_k + \Gamma_{LLR})$$

$$\text{If } \Psi(Y_k + \Gamma_{LLR}^+) > VR_{new}$$

$$\text{– Update } Y_{k+1} = Y_k + \Gamma_{LLR}^+$$

$$\text{– } VR_{new} = \Psi(Y_k + \Gamma_{LLR}^+)$$

$$\text{– } k = k + 1$$

- Exclusion of a classifier (backtrack):

$$\Gamma_{LLR}^- = \operatorname{argmax}_{\Gamma_{LLR} \in Y_k} \Psi(Y_k - \Gamma_{LLR})$$

$$\text{If } \Psi(Y_k - \Gamma_{LLR}^-) > VR_{new}$$

$$\text{– Update } Y_{k-1} = Y_k - \Gamma_{LLR}^-$$

$$\text{– } VR_{new} = \Psi(Y_k - \Gamma_{LLR}^-)$$

$$\text{– } k = k - 1$$

3.6. Fusion of CG Classifiers

We perform log-likelihood ratio (LLR)-based score conversion on the similarity scores of each individual classifier to achieve a higher separability between matching and non-matching classes. If the densities of these two distributions are known, then LLR test achieves the highest VR for a given FAR according to Neyman-Pearson Lemma [4]. Although the real densities are not known, modeling them as two Gaussian distributions with μ_{diff} , Σ_{diff} and μ_{same} , Σ_{same} has proven to work successfully while fusing the classifiers in [8]. In this study, we employ the same idea for conversion of similarity scores as follows:

$$\Gamma_{LLR} = \log \frac{\mathcal{N}(\Gamma; \mu_{same}, \Sigma_{same})}{\mathcal{N}(\Gamma; \mu_{diff}, \Sigma_{diff})} \quad (6)$$

where the parameters, μ_{diff} , Σ_{diff} , μ_{same} , Σ_{same} , are computed from the training set.

After LLR score conversion is applied on each of the CG classifiers, the final classifier is obtained using simple sum fusion of selected CG classifiers. In addition, we combine classifiers trained on different local representations at score-level by learning the weights with PLSR, a statistical technique that generalizes and combines features from the PCA and multiple regression [1]. We trained the PLSR on a randomly generated subset of the training set.

4. Experiments and Results

We evaluated our method on Face Recognition Grand Challenge (FRGC) v2.0 database [15]. The training set consists of 12,776 images from 222 individuals. We used Exp. 4 protocol which is the most challenging experiment because it contains uncontrolled conditions including large illumination changes, out of focus, and partial occlusions. The gallery and probe set of Exp. 4 consist of a single still image per subject. It contains 8,014 uncontrolled query images and 16,028 controlled target images from 466 subjects. The performance measure is the VR at 0.1% FAR. The performances are reported by three Receiving Operator Characteristic (ROC) curves which correspond to three different time gaps. ROC I corresponds to the images collected within a semester, ROC II within a year, and ROC III between semesters.

FRGC dataset provides labels of salient facial features such as eye centers and mouth corners. We use the provided eye labels to align the face images. More specifically, we align the face images into 128×160 pixels with the eye distance being 72 pixels. After 2D CG filtering, the resulting features are spatially partitioned into 20 non-overlapping patches of 32×32 pixels. Since we use 5 scales and 16 orientations, the dimensionality of each local block is $32 \times 32 \times 5 \times 16 = 81,920$ (if $c \neq 0$), which is very high compared to the number of subjects (222) in the training set. So each local block is down-sampled by averaging the magnitude values in an 8×8 grid, and the dimensionality is reduced to 1280 ($= 4 \times 4 \times 5 \times 16$). The resulting CG features are Z-normalized before the subspace analysis.

4.1. Results on Individual CG Classifiers

As ROC III evaluates the matching with large time gap (between semesters), we compared the ROC III performance in the later experiments. The performances of 20 individual CG classifiers are shown in Fig. 5. The best VR, 91.05%, is achieved by a conventional wavelet, $\Gamma_{(c=0.0, \sigma=\pi)}$, while the worst VR, 81.36%, is achieved by a curvature wavelet, $\Gamma_{(c=0.2, \sigma=1.5\pi)}$. The performances of individual CG classifiers are comparable with most of the existing systems. For example, the best individual CG classifier shows a better performance than the ‘‘T-shaped ECG Classifier Bunch’’ (90.36%) [8], therefore, CG features with local face modeling outperform the holistic CG features.

4.2. Results for Fusion of CG Classifiers

After LLR-based score conversion, which can improve the performances of individual CG classifiers up to 0.4%, is applied on each CG classifier, we select a set of them using SFFS-based selection mechanism and fuse their similarity scores with LLR fusion. Fig. 6 shows the effects of classifier selection and fusion. When there is no selection, that

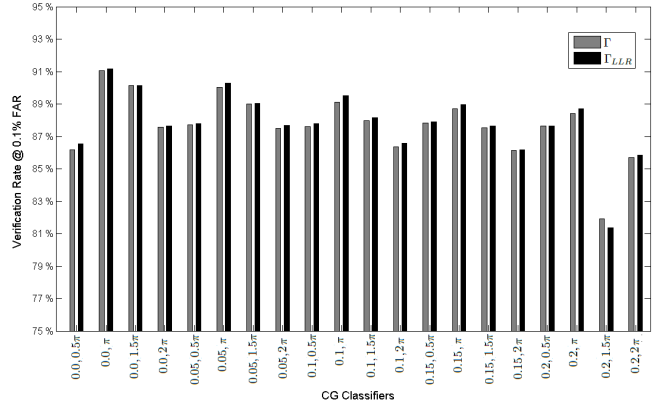


Figure 5. The performances of individual CG classifiers and LLR-based score conversion applied CG classifiers at Exp. 4.

is fusion of all 20 CG classifiers, we achieved 92.85% with LLR fusion, 92.52% with simple sum fusion. When empirical selection, which is selecting the classifiers with the highest VRs, is used, the best VRs are 93.15% with LLR fusion of 15 classifiers and 92.87% simple sum fusion of 4 classifiers. On the other hand, the best fusion performance is achieved by SFFS-based classifier selection algorithm, and we achieved 93.46% by LLR fusion of 8 classifiers.

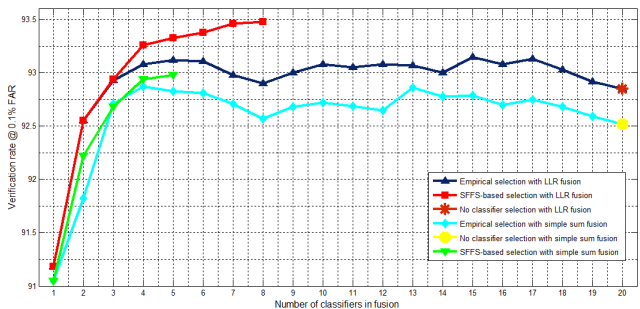


Figure 6. Comparison of classifier selection and fusion methods.

We performed additional experiments to see whether a different face representation provides complementary information for verifying faces. We used another local face representation based on discrete cosine transformation (DCT) [5] as extra evidence. We combined the similarity matrix obtained from this representation with our global CG classifier on score-level by performing weighted sum fusion. Weights of final CG classifier and DCT classifier are learned by PLSR analysis performed on the training data. In this way, the VR is improved further to 94.16%.

4.3. Comparison with Previous Work

Table 4 summarizes the performances of previous studies under the same experiment. Earlier, holistic and homo-

geneous approaches [9], [10] were proposed and their VRs are 74%-76%. Then, researchers tended to combine different types of features and classifiers to improve the results. [19] introduced a method which utilizes DCT and Gabor features, and [18] combined holistic Fourier and local Gabor features. The best performances in the literature, 92.4%-92.5%, were achieved by [12] and [6]. [12] combined 17 classifiers whose features were extracted by DCT, LBP, and Gabor, separately. On the other hand, [6] combined local Gabor classifiers obtained from different image resolutions and a DCT classifier. In addition, [8] proposed to boost holistic CG features on LR images with AdaBoost.

Method	Features	ROCIII(%)
[9]-2006	Holistic Fourier	74.33
[10]-2006	Holistic Gabor with KFA	76
[19]-2007	Holistic Gabor+LBP	83.6
[18]-2009	Holistic Fourier+Local Gabor	89
[12]-2009	Local Gabor+LBP+DCT	92.4
[6]-2010	Multi resolution Gabor+DCT	92.5
[8]-2011	Holistic ECG	90.36
Proposed	Local CG	93.46
Proposed	Local CG+Local DCT	94.16

Table 1. Performance comparison on FRGC v2.0 Exp.4 database.

Hence, our method outperforms all of the previous studies on the same dataset reported in the literature.

5. Conclusion

This paper presents a local appearance-based face verification system that focuses on selecting and combining multiple Gabor classifiers that are trained on different scales and local regions. Local facial appearance is modeled with CG features in addition to conventional Gabor features. To acquire complementary information in different scales, and curvatures, we used different curvature ratios and Gaussian sizes in feature extraction. Our final classifier is obtained by combining selected classifiers using SFFS-based classifier selection mechanism. In addition, we combine classifiers trained on different local representations at score-level by learning the weights with PLSR. The proposed system was evaluated on FRGC v2.0 Exp. 4 data set. We achieved 94.16% VR @0.1% FAR, which is the best result reported in the literature.

References

[1] H. Abdi. Partial least squares regression (pls-regression). *Encyclopedia for research methods for the social sciences. Thousand Oaks (CA): Sage*, pages 725–728, 2003.

[2] M. Bartlett, J. Movellan, and T. Sejnowski. Face recognition by independent component analysis. *IEEE Trans. on Neural Networks*, 13:1450–1464, 2002.

[3] P. Belhumeur, J. Hespanha, and D. Kriegman. Eigenfaces vs. fisherfaces: Recognition using class specific linear projection. *IEEE Trans. on PAMI*, 19(7):711–720, 1997.

[4] T. M. Cover and J. A. Thomas. Elements of information theory. *New York: Wiley*, 1991.

[5] H. K. Ekenel and R. Stiefelhagen. Analysis of local appearance-based face recognition: Effects of feature selection and feature normalization. *Proc. of IEEE CVPR Biometrics Workshop*, 2006.

[6] H. Gao, H. K. Ekenel, M. Fischer, and R. Stiefelhagen. Multi-resolution local appearance-based face verification. *Proc. Int. Conf. on Pattern Rec.*, pages 1501–1504, 2010.

[7] Z. Hafed and M. Levine. Face recognition using the discrete cosine transform. *Int. J. Comput. Vis.*, 43(3):167–188, 2001.

[8] W. Hwang, X. Huang, K. Noh, and J. Kim. Face recognition system using extended curvature gabor classifier bunch for low-resolution face image. *Proc. of IEEE CVPR Workshop*, pages 15–22, 2011.

[9] W. Hwang, G. Park, J. Lee, and S. Kee. Multiple face model of hybrid fourier feature for large face image set. *Proc. IEEE Int. Conf. CVPR*, pages 1574–1581, 2006.

[10] V. Kumar, M. Savvides, and C. Xie. Capitalize on dimensionality increasing techniques for improving face recognition grand challenge performance. *IEEE Trans. Pattern Anal. Mach. Intell.*, 28(5):725–737, 2006.

[11] J. Lai, P. Yuen, and G. Feng. Face rec. using holistic fourier invariant features. *Pattern Recognit.*, 34(1):95–109, 2001.

[12] Z. Liu and C. Liu. Robust face recognition using color information. *Advances in Biometrics*, 5558:122–131, 2009.

[13] T. Mandal and Q. M. J. Wu. Face recognition using curvelet based pca. *Int. Conf. on Pattern Recognition*, 2008.

[14] G. Peters, N. Kruger, and C. von der Malsburg. Learning object representations by clustering banana wavelet responses. *Proc. 1st Int. Workshop Statistical Techniques in Pattern Recognition*, pages 113–118, 1997.

[15] P. J. Phillips, P. Flynn, T. Scruggs, K. Bowyer, J. Chang, K. Hoffman, J. Marques, J. Min, and W. Worek. Overview of the face recognition grand challenge. *Proc. IEEE Int. Conf. CVPR*, 1:947–954, 2005.

[16] P. Pudil, J. Novovicov, and J. Kittler. Floating search methods in feature selection. *Pattern Recognition Letters*, 15:1119–1125, 1994.

[17] M. Savvides, J. Heo, R. Abiantun, C. Xie, and B. V. K. V. Kumar. Class dependent kernel discrete cosine transform features for enhanced holistic face rec. in frgc-ii. *Proc. Int. Conf. Acoustics, Speech and Signal Proc.*, 2:185–188, 2006.

[18] Y. Su, S. Shan, X. Chen, and W. Gao. Hierarchical ensemble of global and local classifiers for face recognition. *IEEE Trans. on Image Processing*, 18(8):1885–1896, 2009.

[19] X. Tan and B. Triggs. Fusing gabor and lbp feature set for kernel-based face rec. *IEEE Int. Workshop on Analysis and Modeling of Face and Gestures*, pages 235–249, 2007.

[20] M. Turk and A. Pentland. Eigenfaces for recognition. *Journal of Cognitive Neuroscience*, 3(1):71–86, 1991.

[21] L. Wiskott, J.-M. Fellous, N. Kruger, and C. von der Malsburg. Face rec. by elastic bunch graph matching. *IEEE Trans. on PAMI*, 19(7):775–779, 1997.