Improving Facial Recognition with Heterogeneous Set of Features

K. R. Singh[1], M. A. Zaveri[2], M.M. Raghuwanshi[3]
[1]Computer Technology Department, Y.C.C.E, Nagpur, 441110, India.
[2]Computer Engineering Department, S.V .National Institute of Technology, Surat, 329507, India.
[3] Rajiv Gandhi College of Engineering and Research, Nagpur, 441110, India.

Abstract - *In this paper we present a feature based fusion approach to face recognition under varying poses. We fuse feature sets extracted from three different face representation techniques, Four Patch Local Binary Pattern (FPLBP), Discrete Cosine Transform (DCT) and 2D log Gabor transform. The features extracted for each of the face representation techniques are higher in dimension. Therefore before fusion of three feature sets, we apply singular value decomposition (SVDs) on every set of feature. Reduced feature sets are then combined for face recognition. Experimentally we show that our approach is efficient in recognising the faces under varying poses, scaling and illumination.*

Keywords: *Feature fusion, DCT, FPLBP, 2D log Gabor.*

1 Introduction

Different face representation techniques have been proposed since last so many decades. Few of them are Gabor transforms [1], Discrete Cosine transform (DCT) [2,3] or Four Phase Local Binary Patterns (FPLBP) [4], etc.. No doubt these techniques can be used to represent the face as a single feature set very efficiently. However, certain challenges due to uncontrolled conditions such as pose, illumination, age, scaling, expression and occlusion still exist with face recognition system [5]. One way to conquer these challenges is through well-built feature sets. This can be accomplished through fusion of different feature sets. Fusion can take place at both the feature and the decision level but we mainly focus on the feature based fusion [4]. From feature level fusion point of view, we propose multilevel fusion of features for representation of face. Here we fuse the features from Discrete Cosine Transform (DCT) [3], 2Dlog Gabor [6] and Four Phase Linear Binary Pattern (FPLBP) [4, 7]. We combine the features at different levels to bring the robustness in recognizing the faces under uncontrolled conditions. The reason for preferring DCT is that it converts high-dimensional face into low-dimensional features in which more significant facial features are maintained. The energy compression nature of DCT makes it to deal with scaling problem in face recognition. On the contrary, the 2D log Gabor transformed face images exhibit strong characteristics of spatial frequency and polar angle selectivity, which produce salient features at various orientations which stand firm to pose variations. Therefore features at all polar angles for every spatial frequency are fuse together to form a single feature vector. Next, FPLBP, is resistant to lighting changes [4], Therefore, FPLBP is a good choice for fusing along with DCT and 2D log Gabor. All the three feature sets are high in dimension so it is beneficial to use SVDs to reduce the dimension of individual feature set prior to fusion. Simple concatenation method is then applied to the reduced feature sets. Fused feature set is then used for recognition using nearest neighbor classifier. We assess the proposed approach on quite a lot of challenging face databases including ORL, Georgia, Head Pose Image Database and CMU-PIE with promising outcome

In the paper Section 2 gives the fusion method of FPLBP, DCT and 2D log Gabor. Section 3 explains the experimentation for parameter selection and evaluation followed by conclusion.

2 Fusing FPLBP, DCT and 2Dlog Gabor

In this section, specifically we discuss the fusing methodology of feature sets from thee different face representation techniques. First to produce FPLBP feature set, FPLBP [4] is applied on a face image. It returns two parameters; one FPLBP code and other descriptor matrix. In our case we have only considered the descriptor matrix as FPLBP features values.

$$FPLBP^{d1} = (F_{B1}, F_{B2}, \dots \dots \dots F_{BN}) \qquad (1)$$

where, F_B is the block applied on each pixel for feature extraction. All the extracted blocks are combined to form a $FPLBP$ feature vector. To neutralize the high dimension of resultant feature set we apply SVDs for feature reduction. Next, the facial image $I(x, y)$ is represented with DCT. The observation sequence is obtained by sliding a square fixed size window over the face image, in a raster scan fashion, with a predefined overlap. It is worth noting that most of the transformed coefficients have very small values and only a few coefficients have higher magnitudes on the lower side. Therefore, from the obtained coefficients few of coefficients (say 15 represented as M) are retained. This M coefficients determines the dimensionality of the observation. To retain these M significant coefficients we used zonal coding along with threshold coding. For an image of size 64 × 64, with a sliding window of 16 × 16 with 75% overlapping results in (N= 169) blocks (B). The value of each block will be observed as observation vector of size 1 X M coefficients. The concatenation of all these observation vectors results in a final DCT face matrix DCT^{d2} as shown in Equation 2; of size N × M (d=169×15=2535) coefficients. To this high dimensional matrix we apply SVDs, which returns a vector of size 1×15.

$$DCT^{d2} = (B^1,\ B^2, \dots \dots \dots B^N) \tag{2}$$

where, d_2 is the dimensionality of the DCT feature vector.

At last, an image $I(x, y)$ is convolved with family of Gabor filters at spatial frequency r and polar angle θ.

$$G_{r,\theta}(x, y) = \left(I * \varphi_{r,\theta}\right)(x, y) \tag{3}$$

where, $G_{r,\theta}(x, y)$ denotes the convolution result corresponding to the Gabor filter at spatial frequency r and polar angle θ. In log polar Gabor transform the radial distance represents the spatial frequency and the polar angle represents the orientation. As a result, image $I(x, y)$ can be represented by a set of Gabor coefficients $\{G_{r,\theta}(x, y),\ r = 0, \dots 4;\ \ \theta = 0, \dots \dots 7\}$. The magnitudes of 2D log Gabor $\{G_{r,\theta}(x, y)\}$ for all the polar angles θ at each spatial frequency r are concatenated. First the magnitude of each $\{G_{r,\theta}(x, y)\}$ at r=0 and $\theta = 0$ are reduced with SVDs, and turned to a vector $G^{0,0}_{r,\theta}$. The concatenation of these eight vectors forms a discriminative Gabor feature vector $G^{0}_{r,\theta}$ at r=0.

$$G^0_{r,\theta} = (G^{0,0}_{r,\theta}, G^{0,1}_{r,\theta}, G^{0,2}_{r,\theta}, G^{0,3}_{r,\theta}, G^{0,4}_{r,\theta}, G^{0,5}_{r,\theta}, G^{0,6}_{r,\theta}, G^{0,7}_{r,\theta}) \tag{4}$$

Similarly, final 2D log Gabor transform feature vector set G^{d3} for a single image is thus derived by concatenating all the 2D log Gabor feature vectors that encompasses all the Gabor coefficients of the image as shown in Equation (5);

$$G^{d3} = (G^0_{r,\theta}, G^1_{r,\theta},, G^2_{r,\theta \dots\dots\dots\dots\dots\dots}, G^4_{r,\theta}) \tag{5}$$

where, d_3 is the dimension of 2D log Gabor feature set. Now the feature vectors represented through three different face descriptors are combined through concatenation approach. Here, prior to concatenation only, feature reduction has been performed in order to reduce the counter effect of high feature dimension. Thus the dimension of the resulting feature vector is equal the sum of the dimension of the feature vectors of all the three face representation techniques.

$$F^d = (FPLBP^{d1}, DCT^{d2}, G^{d3}) \tag{6}$$

where, $d = d_1 + d_2 + d_3$ is the dimension of fused feature vector for a facial image. Once the fused feature is formed we can use this information to recognize that particular face image in à database.

3 Experiment Setup

In this section, we evaluate our approach on ORL [8], Head Pose Image Database [10], Georgian [9] and CMU-PIE [11]. The details of each database can be found in [8, 9, 10, 11]. First, we briefly describe the parameter selection procedure for each type of face descriptors. Then with the selected parameters we evaluate our proposed approach under different experimental testing sets for various face databases.

3.1 Parameter selection

In our first experiment, we intend to find the most excellent 2D log Gabor filter parameters for our experiment. We conducted the experiment for each possible combination of spatial frequency and polar angle. The graph in Figure 1 shows the performance of the face recognition system for few possible combination of spatial frequencies and polar angles for different face databases. It can be observed from the Figure 1 that the performance increases with the number of spatial frequency and number of polar angle. However, when the number of spatial frequency reaches five, the recognition rate becomes stable. Since most of the valuable information in face images is contained within a limited frequency band, the inclusion of more scales will result in redundant information and thus reduce system performance. Five spatial frequencies and eight polar angles appear to have achieved good performance in our experiments. Therefore, we choose to use 2D log Gabor filters of five spatial frequencies and eight polar angles for further experiments.

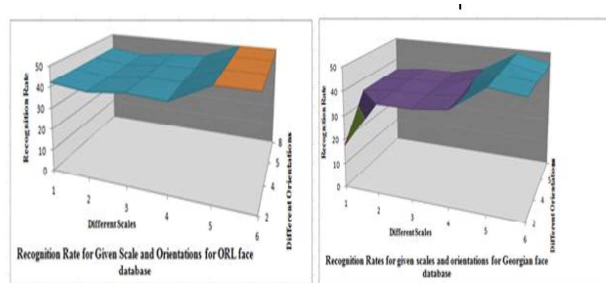

Figure 1: Recognition rates at different spatial frequency (scales) and polar angle (orientations).

In the next experiment of parameter selection, we aim to find the effect of sampling window size and percentage of overlapping on the performance of the face recognition system. We conducted the experiment for sampling window of size 8 X 8 and 16 X 16. For each sampling window size we repeated the experiment for 50% overlapping and 75% overlapping. With the aid of the results summarized in Table I, appropriate parameters such as the sampling window size of 8 X 8 and 75 % percentage of overlapping for further procedure has been chosen. Another parameter that plays an important role is the size of an image. Since the sampling window is in multiples of two, size of image has also to be in multiples of two. For this experimentation sampling window is chosen as 8 X 8 with 75% overlapping. The results of this experiment along with training time and testing time are summarized in Table II only for two face databases due to lack of space. It is to be noted that highest recognition rates can be achieved for 64 X 64 image size. Further, if image size is increased, recognition rate starts deteriorating with enormous increase in training and testing time. Based on the above results appropriate image size of 64 X 64 for DCT is chosen.

In the last parameter selection experiment for FPLBP, an important consideration is again the image size. In this experiment recognition rate for different image sizes is summarized in Table III. Here – indicates that results were poor for this image sizes so they are not listed in Table III. It must be noted that recognition rate increases with increase in image size. This is because with higher image sizes more greater numbers of descriptors are returned which in turn gives better number of FPLBP coefficients. Image chosen to be 128×128 pixel size for further experimentation.

Table I: Comparative recognition rate for different size of sampling window and overlapping (DCT).

Overlapping	Recognition Results		
	Databases	Sampling Window	
		8X8	16X16
50	ORL	48.88%	44%
	CMU-PIE	40.23%	30.90%
	Georgian	58.80%	54.21%
	Head Pose	60%	59.00%
75	ORL	**62%**	50%
	CMU-PIE	**58.90%**	50.23%
	Georgian	**63.02%**	60.14%
	Head Pose	**61.90%**	58.89%

Table II: Comparative recognition rate for different image sizes (DCT)

Face Database	Image Size	Recognition Rate (%)	Training Time(sec)	Testing Time(sec)
Georgian	32	17.77	1.24	0.63
	64	**20**	2.19	1.46
	128	15.55	5.90	5.36
	256	13.33	24.6	23.9
CMU_PIE	32	18.88	1.044	2.22
	64	**19.66**	1.55	2.82
	128	12.22	4.12	10.2
	256	15.55	16.57	47.61

Table III: Comparative recognition rate (%) for different image sizes (ORL database)

→ X ↓ Y	16	32	42	64	128
16	-	-	40.3	48	66.6
32	-	58.6	60.0	66.6	80
48	30.6	74.6	76	82.67	80
64	52.3	76	80.2	89.3	90.6
128	70.0	90.6	88.3	89.3	**93.3**

3.2 Experimental Procedure

For experimentation each database is partitioned into training and testing sets. In all the experiments reported in this work, individual training sets for each type of face database have been formed from five samples for each subject. Before applying the feature extraction other pre-processing steps are applied such as face is located in the image to remove it from the background and resized to uniform size. Then, face images are normalized with histogram equalization to equalize the illumination problem.

3.3 Evaluation Results

For evaluation, we selected 200 images for training and remaining 200 images of 40 subjects (10 poses within $\pm 20^{\circ}$ in yaw per subject) for testing from ORL database. For Georgian face database 150 images were used for training. Testing was performed on two separate sets frontal and poses variation. From the Head Pose Image database, for each tilt orientation ($0^{\circ}, \pm 15,^{\circ} \pm 30^{\circ}$) we considered 9 horizontal pose variations ($0^{\circ}, \pm 15,^{\circ} \pm 30,^{\circ} \pm 45,^{\circ} \pm 60^{\circ}$). Thus testing was performed on total 1350 images from 45 separate tests. For instance, each set named as; TS+30-45; represents testing set consisting of images with +30 tilt orientation and – 45^{0} variations in yaw. For CMU-PIE database we evaluated the approach on five different testing sets with varying number of images as revealed in Table IV.

Based on the parameter selected from parameter selection section, we evaluated the proposed approach on different testing sets. With aid of the results summarized in Table IV, it can be observed that for a perfect database such as ORL with lesser variations in pose, single feature sets also suffice and gives better recognition rate. Whereas, for frontal set of Georgian face database, 2D log Gabor and DCT has performed badly except FPLBP features. This is due to the fact that Georgian face database contains subjects with huge variation in skin tone and FPLBP is resistant to this variation. For the same frontal experimental set our approach gives 96.0% of recognition. 2D log Gabor features contain more discriminant information regarding orientations and are thus more robust against variations in pose and expressions.

However, due to variation in skin tone even 2D log Gabor shows deprived result for pose variation set of Georgian face database. For CMU-PIE database we evaluated on 5 separate test sets. For scaling variation, expression variation, pose variation and illumination variation our feature fusion have shown better result compared to the performance shown by single feature set for these testing sets. Although the performance of our approach degrade from 100% to 85% still it shows consistency compared to the performances of single feature sets.

4 Conclusions

In this paper, we have proposed feature based fusion of FPLBP, DCT and 2D log Gabor for face recognition. The design of each face representation technique has also been discussed and experimentally required parameters are tuned for face recognition. Experiments have been conducted on ample number of facial images having variations in pose, scale, illumination and expression. The proposed approach is evaluated using the ORL, Georgian, Head Pose Image Face

Database and CMU-PIE databases. The proposed approach shows significantly improved recognition rate than the single feature sets. Though the performance of our approach has been extensively tested and evaluated various databases, we are also working with other feature reduction techniques and recognizers.

5 References

[1] J. P. Jones and L. A. Palmer, "An evaluation of the two-dimensional Gabor filter model of simple receptive fields in cat striate cortex," *J. Neurophysiol.*, vol. 58, no. 6, pp. 1233–1258, 1987.17

[2] J. Lai, P. C. Yuen, and G. Feng, "Face recognition using holistic Fourier invariant features," *Pattern Recognit.*, vol. 34, no. 1, pp. 95–109, 2001.

[3] Vinayadatt V. Kohir and U. B. Desai, 'Face Recognition Using A DCT-HMM Approach', *IIT Bombay*, 2008

[4] L.Wolf, T.Hassner, and Y.Taigman1, "Descriptor Based Methods in the Wild," Faces in Real-Life Images workshop at the European Conference on Computer Vision (ECCV), Oct 2008

[5] W. Zhao, R. Chellappa, P. J. Phillips, and A. Rosenfeld, "Face recognition: A literature survey," *ACM Computing Survey*, 34(4):399–485, 2003.

[6] M. Lades, J. C. Vorbruggen, J. Buhmann, J. Lange, C. von der Malsburg, R. P. Wurtz, and W. Konen, "Distortion invariant object recognition in the dynamic link architecture," *IEEE Trans. Comput.*, 42(3):300–311, 1993.

[7] T. Ojala, M. Pietikainen, and D. Harwood, "A comparative study of texture measures with classification based on feature distributions. Pattern Recognition," Vol: 29, Issue: 1, pp: 51-59, 1996.

[8] ORL web site: *http://www.camorl.co.uk*.

[9] Georgia Tech Face Database, ftp://ftp.ee.gatech.edu/pub/users/hayes/facedb/.

[10] N.Gourier, D.Hall and J.L.Crowley; "Estimating Face orientation from Robust Detection of Salient Facial Features", *In Proceedings of pointing 2004, international Workshop on Visual Observation of Deictic Gestures*, Cambridge, UK, 2004.

[11] T. Sim, S. Baker, and M. Bsat, "The CMU-PIE Database," *Proceedings of the IEEE International Conference on Automatic Face and Gesture Recognition*, May 2002.

Table IV: Comparative performance in (%) of proposed approach on different databases.

Approaches / Databases	Number of Test Images	2d log Gabor +Knn	DCT +Knn	FPL BP +Knn	Our Approach
ORL	200	100	100	100	100
Georgian	Frontal (260)	45	48.33	78.33	96.0
	Pose Variation(600)	40.83	56.67	40.98	85.8
Head Pose Image Face Database	Frontal+ Pose variation (1350)	80.00	45.00	65.00	83.4
CMU-PIE	Frontal (600)	98.34	100	99	100
	Pose Variation(500)	89.34	66.00	65.9	93.0
	Scaling Variation(200)	73.00	79.00	67.0	85.2
	Expression Variation (200)	63.00	72.00	79.00	89.7
	Illumination Variation (400)	56.00	69.00	80.00	95.0