

Face Recognition using DCT based Energy Discriminant Mask

Vikas Maheshkar
Division of Information technology
NSIT
New Delhi, India

Sushila Maheshkar
Department of Computer Science and Engg
IIT (Indian School of Mines)
Dhanbad, India

ABSTRACT

It has been observed that the variations among the images of the same face due to illumination and viewing direction are almost always larger than image variations. One person, with the same facial expression, can appear strikingly different when light source direction and viewpoint vary. These variations are emphasized by additional factors such as facial expressions, perspiration, hair style, cosmetics, and even changes due to aging. The proposed Face recognition technique is based on Energy discriminant mask obtained by thresholding DCT coefficients in low, mid and high frequency regions. The proposed approach analyzes all images of a database to know the discrimination ability of individual DCT coefficient and generates a database specific DCT mask. High recognition rate can be achieved by using the coefficients that have maximum discrimination power. To benchmark proposed techniques standard ORL and YALE face databases are used.

Keywords

Biometrics, Face recognition, Discriminant, DCT, Mask, Energy.

1. INTRODUCTION

The recent progress of digital multimedia technologies has offered many facilities in transmission of information but at the same time it increases the security concerns. As the level of security breaches and transaction fraud increases, the need of highly secure identification and personal verification technologies increases. Biometric technologies are becoming the foundation of highly secure identification and personal verification solutions. Biometric-based authentication applications include workstation, network, domain access, single sign-on, application logon, Data protection, remote access to resources, transaction security and Web security [1], [12]. A face recognition system is a computer application for automatically identifying or verifying a person from a digital image or a video frame captured using video source. The common issues of face recognition remain to be unsolved, since many of the work done can successfully recognize faces, only when images are obtained under prescribed conditions. The performance of the system will degrade abruptly when face images are captured under varying pose, lighting, with accessories and expressions [19]. Face recognition must be robust enough, so that the image variations caused by the illumination conditions, viewing conditions or poses, different facial expression, ageing, and disguises such as slight cut, glasses or makeup do not affect the recognition. Automatic face recognition was developed in early 1960s. Over the last 20 years, several different techniques have been proposed for computer recognition of human faces. Zhao et al. gives an excellent survey on many approaches for holistic feature [19]. The machine recognition systems have reached a certain level of maturity but their success is limited by the conditions

imposed by many real applications. For example, recognition of face images acquired in an outdoor environment with changes in illumination and or pose remains a largely unsolved problem [11]. In other words, current systems are still far away from the capability of the human perception system [8], [10]. Two important issues in practical face recognition systems are identified: the illumination problem and the pose problem [14], [5]. Kanade extracted geometric features for face recognition. Facial feature points such as eye corners, nose, mouth, and chin counter are located [11]. The template matching strategy has the advantage of speed with restriction of being sensitive to translation, size, illumination and orientation [6]. Jia and Nixon improved feature definition and extraction for frontal view image, by calculating measurements of various face regions [7]. Zhou et al. captured facial local features using discrete cosine transform [20]. Mayer et al. evaluated the color information and location based on data pixel features for face recognition [9]. Dabbaghchian et al. used DCT to extract features for face recognition [3]. Yin et al. used DCT and two dimensional linear discriminant analysis (2DLDA) for face recognition [18]. Choi et al. novel feature extraction method for face recognition is based on DCT, energy probability (EP), and linear discriminant analysis (LDA) [2]. Dandpat and Meher automatic face recognition system used DCT for feature extraction and normalization techniques are invoked to increase the robustness and to capture the variations in facial geometry and illumination [15].

The proposed approach in this paper uses DCT coefficients since they do not have the same ability to discriminate various classes. Some coefficients named as discriminant power coefficients, separate classes better than others. The proposed approach statistically analyzes all images of a database and retains more important DCT coefficients according to its Energy discrimination power from low, mid and high frequency regions. This analysis is named as Energy Discrimination Power Analysis (EDPA). Energy discrimination power analysis generates a database specific DCT mask. This approach is able to find the best discriminant coefficients for particular database. It is found that feature dimensionality reduction does not significantly affect the discrimination power.

The rest of the paper is organized as follows. The proposed technique is explained in section 2 followed by experimental results and conclusion in section 3.

2. PROPOSED TECHNIQUE

The coefficient selection strategy of the DCT is an important part of the feature extraction process. In most of the approaches which uses the DCT, zigzag scan or zonal masking is preferred. These conventional approaches are not necessarily efficient in generating the discriminant coefficients. The proposed Energy Discrimination Power

Analysis (EDPA) is based on Energy Discriminant Mask obtained by thresholding DCT coefficients in Low, Mid and High frequency regions for face Recognition. EDPA coefficients depend on large variation between the classes and small variation within the classes and can be estimated by the division of the between-class variance to the within-class variance. EDPA limit the number of DCT coefficients by masking and generating a database specific DCT mask. The proposed approach is as follows:

Step 1: Compute the DCT of all the images in the database. DCT coefficient matrix for first image of size $N \times N$ is represented by:

$$X = \begin{bmatrix} x_{1,1} & x_{1,2} & \dots & x_{1,N} \\ x_{2,1} & x_{2,2} & \dots & x_{2,N} \\ \cdot & \cdot & \cdot & \cdot \\ x_{N,1} & x_{N,2} & \dots & x_{N,N} \end{bmatrix}_{N \times N}$$

Step 2: From the DCT coefficient for all the images in the database construct the other matrix $A_{i,j}$ as follows:

$$A_{i,j} = \begin{bmatrix} x_{i,j}(1,1) & x_{i,j}(1,2) & \dots & x_{i,j}(1,C) \\ x_{i,j}(2,1) & x_{i,j}(2,2) & \dots & x_{i,j}(2,C) \\ \cdot & \cdot & \cdot & \cdot \\ x_{i,j}(S,1) & x_{i,j}(S,2) & \dots & x_{i,j}(S,C) \end{bmatrix}_{S \times C}$$

$i = 1, 2, \dots, N; j = 1, 2, \dots, N$

Where, C is number of classes and S is the number of images for each class. Total $C \times S$ images are present in the database.

$x_{i,j}(l,m)$ is the $(i,j)^{\text{th}}$ DCT Coefficient of the l^{th} face in the m^{th} class.

Step 3: Calculate the mean value at $(i, j)^{\text{th}}$ position for each class:

$$M_{i,j}^C = \frac{1}{S} \sum_{s=1}^S A_{i,j}(s, c) \quad c = 1, 2, \dots, C$$

Step 4: Calculate variance at $(i, j)^{\text{th}}$ position for each class:

$$V_{i,j}^C = \sum_{s=1}^S (A_{i,j}(s, c) - M_{i,j}^C)^2 \quad c = 1, 2, \dots, C$$

Step 5: Average the variance of all the classes with respect to position (i, j) :

$$V_{i,j}^W = \frac{1}{C} \sum_{c=1}^C V_{i,j}^C$$

Step 6: Calculate the mean of the database samples

$$M_{i,j} = \frac{1}{S \times C} \sum_{c=1}^C \sum_{s=1}^S A_{i,j}(s, c)$$

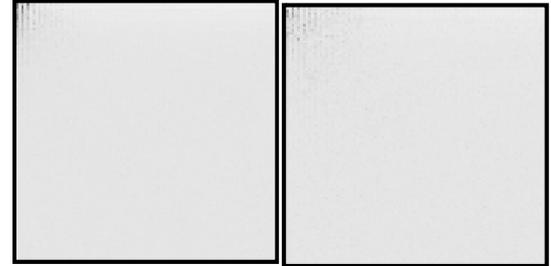
Step 7: Calculate the variance of the database samples

$$V_{i,j}^B = \sum_{c=1}^C \sum_{s=1}^S (A_{i,j}(s, c) - M_{i,j})^2$$

Step 8: Estimate the DP for location (i, j) :

$$DP(i, j) = \frac{V_{i,j}^B}{V_{i,j}^W}, \quad 1 \leq i \leq N, \quad 1 \leq j \leq N$$

Thus, the Discrimination Power(DP) matrix is obtained by the ratio of between-class variance and within-class variance. Figure 2.1 (a) and Figure 2.1 (b) shows the DP matrix for ORL and YALE Database, respectively.



(a) ORL dataset (b) YALE dataset
Fig 2.1: Discrimination Power matrix

The low-frequency coefficients have large discrimination power whereas high frequency coefficients have small power of discrimination.

Step 9: Constitute a zigzag scan sequence of DP matrix giving a vector DPs of low, mid and high.

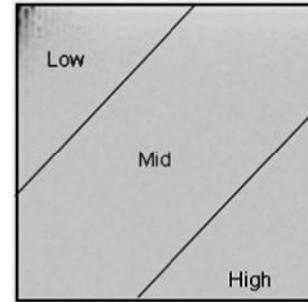


Fig 2.2: Division of DPs matrix into low, mid and high frequency regions size $N^2 \times 1$.

Divide vector DPs equally into three regions as:

$$DP_{LOW}(k) = DP_s(k) \text{ for } 1 \leq k \leq N^2 / 3$$

$$DP_{MID}(k) = DP_s(k - (N^2 / 3)) \text{ for } (N^2 / 3) + 1 \leq k \leq (2N^2 / 3)$$

$$DP_{HIGH}(k - (2N^2 / 3)) = DP_s(k) \text{ for } (2N^2 / 3) + 1 \leq k \leq N^2$$

DP_{LOW} contains the low frequency information and is related with the smooth regions. DP_{MID} contains the middle frequency information and have the basic structure. DP_{HIGH} contains high frequency information related with the edges regions. Figure 2.2 shows the division of DPs into low, mid and high frequency regions.

Step 10: Energy of each region (Low, Mid, High) from DP matrix is calculated as:

$$DPE_{LOW} = \sum_{k=1}^{N^2/3} |DP_s(k)|^2$$

$$DPE_{MID} = \sum_{k=N^2/3+1}^{2N^2/3} |DP_s(k)|^2$$

$$DPE_{HIGH} = \sum_{k=2N^2/3+1}^{N^2} |DP_s(k)|^2$$

Step 11: Energy probability (EP) of each region (Low, Mid, High) is calculated as:

$$EP_{LOW}(k) = \frac{|DP_s(k)_{LOW}|^2}{DPE_{LOW}} \text{ for } 1 \leq k \leq N^2/3$$

$$EP_{MID}(k) = \frac{|DP_s(k)_{MID}|^2}{DPE_{MID}} \text{ for } (N^2/3)+1 \leq k \leq 2N^2/3$$

$$EP_{HIGH}(k) = \frac{|DP_s(k)_{HIGH}|^2}{DPE_{HIGH}} \text{ for } (2N^2/3)+1 \leq k \leq N^2$$

The magnitude of Energy Probability of each region is used as a criterion for selection of discriminant information. The Energy Probability indicates the energy contribution of individual coefficients according to energy in low, mid and high frequency regions.

Step 12: Mean of each region of Energy Probability (Low, Mid, High) is calculated as:

$$MEP_{LOW}(k) = \frac{1}{N^2/3} \sum_{k=1}^{N^2/3} EP_{LOW}(k)$$

$$MEP_{MID}(k) = \frac{1}{N^2/3} \sum_{k=(N^2/3+1)}^{2N^2/3} EP_{MID}(k)$$

$$MEP_{HIGH}(k) = \frac{1}{N^2/3} \sum_{k=(2N^2/3+1)}^{N^2} EP_{HIGH}(k)$$

Step 13: Mean of energy probability (MEP) for each (Low, Mid, High) region is used as thresholds

$$EP_{Low}(k) = \begin{cases} 0 & \text{if } EP_{Low}(k) \leq MEP_{LOW} \text{ for } 1 \leq k \leq N^2/3 \\ 1 & \text{otherwise} \end{cases}$$

$$EP_{Mid}(k) = \begin{cases} 0 & \text{if } EP_{Mid}(k) \leq MEP_{MID} \text{ for } (N^2/3) \leq k \leq (2N^2/3) \\ 1 & \text{otherwise} \end{cases}$$

$$EP_{High}(k) = \begin{cases} 0 & \text{if } EP_{High}(k) \leq MEP_{HIGH} \text{ for } (2N^2/3) \leq k \leq (N^2) \\ 1 & \text{otherwise} \end{cases}$$

To get binary values from EP, the threshold for energy probability in each frequency region is calculated. Values greater than mean of the energy probability (MEP) in each frequency region are set to 1 in the corresponding position of EP. Values less than mean of the energy probability (MEP) in each frequency region are set to 0 in the corresponding position of EP. Merge the binary EP_{Low}, EP_{Mid} and EP_{High} to get EP(k). This procedure will reconstitute the EP.

Step 14: Inverse Zigzag scan of thresholds EP constitute EP mask and is named as EDPA mask as shown in Figure 2.3.

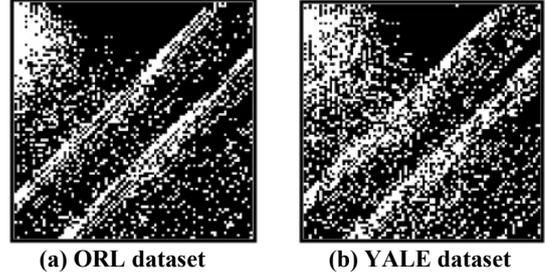


Fig 2.3: EDPA mask

Figure 2.3(a) and Figure 2.3(b) shows EDPA mask for ORL and YALE Database, respectively.

Step 15: This step will determine the discriminant DCT coefficients using EP in Low, Mid and High frequency regions as follows:

- I. Multiply DCT coefficients of each image in the selected training set by the corresponding value in the EDPA mask and construct 1-D feature vectors by zigzag scan of this 2D matrix.
- II. Data reduction is achieved by removing the zero value coefficients from the resulting vector.
- III. Same procedure is adopted for test image to form 1-D vector.

Step 16: Finally, for face recognition, distance between all these training set feature vectors and test image feature vector is calculated. Mahalanobis distance measure is used as the distance measure criteria and the minimum distance measure will give the recognized face [17].

3. EXPERIMENTAL RESULTS AND DISCUSSION

The experimentation of proposed technique is performed on ORL [13] and YALE [4] database. MATLAB is used as simulation tool for experimentation. Results have been analyzed by varying the number of training and test set such that training set has an empty intersection with the test set. The feature dimensionality reduction is achieved by multiplying EDPA mask with the training and test datasets.

Feature dimensionality reduction (FDR) is calculated using the equation given below

$$FDR(\%) = 100 - \left[\frac{(\text{sizeofEDPADCTMask}) \times 100}{\text{sizeoforiginalimage}} \right]$$

The FDR for ORL dataset is found to be 74.40% and for YALE dataset it is 67.36%.

$$FDR(\%) = 100 - \left[\frac{(2166) \times 100}{(8464)} \right] = 74.41$$

$$FDR(\%) = 100 - \left[\frac{(2673) \times 100}{(8464)} \right] = 67.36$$

In order to compare the proposed approach, recognition rate and false recognition rate are computed using equations given below

$$\text{Recognition Rate} = N_c / N_t$$

Where N_c = number of correctly recognized images

N_t = number of test set images

$$\text{False Recognition Rate} = N_m / N_t$$

Where N_m = number of misclassified images

N_t = number of test set image

It is observed that the average recognition accuracy of the proposed technique is 96.22% in case of ORL dataset whereas 96.77% in case of YALE dataset. The highest recognition rate is found to be 99.68% in case of ORL dataset on 8 training images per set whereas 98.64% in case of YALE dataset on 9 training images per set. The results are shown in Table 3.1 and Table 3.2.

It is observed that feature dimensionality reduction does not significantly affect the discrimination power. The false recognition rate is achieved as 0.054 in case of ORL dataset and 0.14 in case of YALE dataset on average. The least false recognition rate is 0.012 on 9 training images in case of ORL dataset and 0.11 on 10 training images in case of YALE dataset. This indicates that the false recognition rate is minimum when more number of training images are taken into consideration. Also, the proposed technique is compared with the existing techniques such as DPA and PM2 with DPA [3]. It is observed that the false recognition rate achieved by proposed technique is less as compared to the proposed technique [3]. This is shown in Table 3.3 and Table 3.4.

The comparative recognition performance plot in terms of false recognition rate of existing techniques with the proposed technique in case of ORL and YALE dataset is shown in Figure 3.1 and Figure 3.2, respectively. It is observed from the plot that the false recognition rate achieved is less as compared to the other existing techniques. As expected, data exhibits a monotonically decreasing false recognition rate for increasing number of training samples per class.

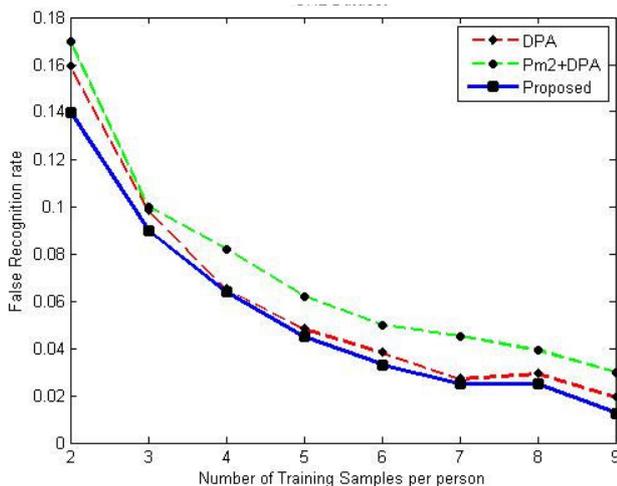


Fig 3.1: Comparative recognition performance of existing techniques with proposed techniques on ORL dataset

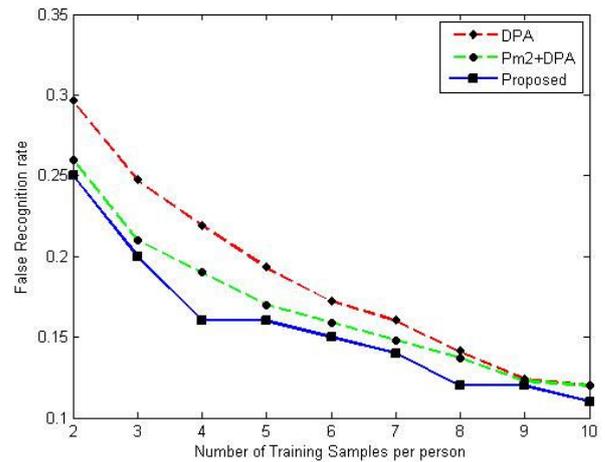


Fig 3.2: Comparative recognition performance of existing techniques with proposed techniques on YALE dataset

The proposed approach shows improvement in recognition accuracy over Eigenface technique [16] and Fisherface technique [1] by 10.22% and 8.15%, respectively. Significant reduction in false recognition rate is also achieved over feature extraction approach using DCT and discrimination power [3]. The comparison of recognition performance of the proposed technique is summarized in Figure 3.3 and Figure 3.4 in case of ORL and YALE dataset, respectively.

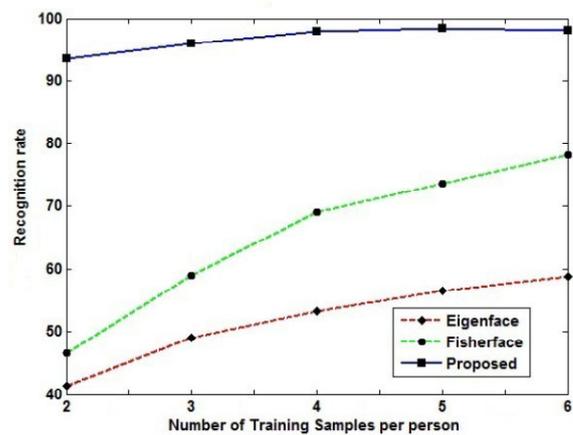


Fig 3.3: Recognition performance comparison of proposed technique for different training sets on ORL dataset

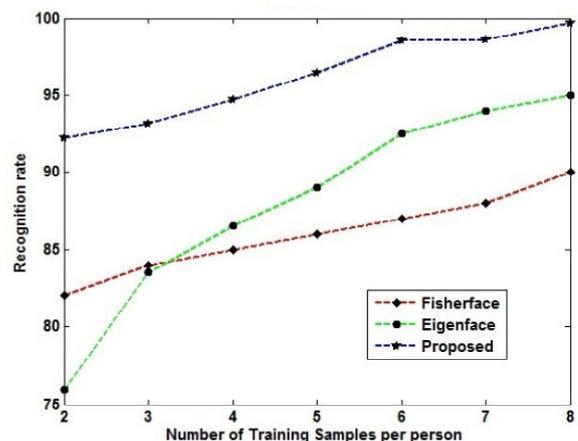


Fig 3.4: Recognition performance comparison of proposed technique for different training sets on YALE dataset

It is observed that recognition is consistently good in terms of recognition rate for all training sets. The training time for ORL database and YALE database is 29.0497 seconds and 10.6249 seconds, respectively. The average testing time for ORL dataset and YALE dataset is 2.5927 seconds and 1.4549 seconds, respectively. Thus, in the proposed technique EDPA mask generated by taking into consideration Low, Mid and High frequency regions for feature selection achieve higher recognition rate.

4. CONCLUSION

In this Paper, Energy Discrimination Power Analysis (EDPA) is proposed using discriminant coefficients of a database for face recognition. EDPA mask is generated using the between class variances and within class variances of DCT coefficients of low, mid and high frequency regions. For dimension reduction, without losing much information, some coefficients are selected from Low, Mid and High frequency regions so as to include discriminant information. To select these coefficients energy probability in each frequency region is calculated. Depending on the thresholds the numbers of discriminant coefficients are selected. An experimental result shows that the proposed EDPA mask is successful in retaining more discriminant coefficients compared to the existing techniques. A comparison of the proposed technique with the existing techniques shows an improvement in recognition rate. Significant dimensionality reduction is achieved using this EDPA mask. YALE database contain images with more illumination variation hence EDPA mask has more number of discriminant coefficients compared to ORL database. Also, it is observed that feature dimensionality reduction does not significantly affect the discrimination power. Another advantage of our approach is its low computational complexity during testing stage. Proposed approach uses DCT for dimension reduction, which means that this approach can be implemented more efficiently than the traditional LDA based methods, especially while running in a large face database. In proposed technique, the coefficients selection strategy from low, mid and high frequency regions, influence an interesting issue for future work.

5. REFERENCES

- [1] Belhumeur, P., Hespanha, J., and Kriegman, D. Eigenfaces Vs. Fisherfaces: Recognition using class specific linear projection. *IEEE transactions on pattern analysis and machine intelligence* 19, 7 (1997), 711–720.
- [2] Choi, J., Chung, Y.-S., Kim, K.-H., and Yoo, J.-H. Face recognition using energy probability in DCT domain. in 2006 IEEE international conference on multimedia and expo (2006), pp. 1549–1552.
- [3] Dabbaghchian, S., Ghaemmaghami, M. P., and Aghagolzadeh, A. Feature extraction using discrete cosine transform and discrimination power analysis with a face recognition technology. *pattern recognition* 43, 4 (2010), 1431–1440
- [4] Georghiades, A. Yale face database. center for computational vision and control at yale university. <http://cvc.yale.edu/projects/yalefaces/yalefaces> (1997).
- [5] Georghiades, A., Belhumeur, P., and Kriegman, D. Illumination cone models for face recognition under variable lighting and pose. *IEEE trans. pattern anal. mach. intelligence* 23, 6 (2001), 643–660.
- [6] Gross, R., and Brajovic, V. An image preprocessing algorithm for illumination invariant face recognition. 4th international conference on audio- and video-based biometric person authentication (berlin, heidelberg, 2003), avbpa'03, springer-verlag, pp. 10–18.
- [7] Jia, X., and Nixon, M. S. Extending the feature vector for automatic face recognition. *IEEE transactions on pattern analysis and machine intelligence* 17, 12 (1995), 1167–1176.
- [8] Kaya, Y., and Kobayashi, K. A basic study on human face recognition. *frontiers of pattern recognition* (1972), 265–290.
- [9] Mayer, C., Wimmer, M., and Radig, B. Adjusted pixel features for robust facial component classification. *Image Vision Comput.* 28, 5 (may 2010), 762–771.
- [10] Norton, E. Identifying the brain's own facial recognition system. 2012
- [11] Sakai, T., Nagao, M., Kanade, T., and University, K. Computer analysis and classification of photographs of human faces. Kyoto University, 1972.
- [12] Samal, A., and Iyengar, P. A. Automatic recognition and analysis of human faces and facial expressions: A survey. *pattern recogn.* 25, 1(jan 1992), 65–77.
- [13] Samaria, F. S., and Harter, A. C. Parameterisation of a stochastic model for human face identification. in proceedings of the second IEEE workshop on applications of computer vision, 1994. (1994), IEEE, pp. 138–142.
- [14] Štruc, V., and Pavešić, N. Performance evaluation of photometric normalization techniques for illumination invariant face recognition. *advances in face image analysis: techniques and technologies* (2010).
- [15] Swarup Ku. and Dandpat, S. M. New technique for DCT-PCA based face recognition. In international conference on electronic systems, national institute of technology, Rourkela, India (2011).
- [16] Turk, M., and Pentland, A. Face recognition using eigenfaces. in IEEE computer society conference on computer vision and pattern recognition, 1991. proceedings CVPR '91. (1991), pp. 586–591.
- [17] Wölfel, M., And Ekenel, H. K. Feature weighted mahalanobis distance: improved robustness for gaussian classifiers. In *13th European Signal Processing Conference* (2005).
- [18] Yin, H., Fu, P., and Qiao, J. Face recognition based on dct and 2dlda. in second international conference on innovative computing, information and control, 2007. ICICIC '07. (2007), pp. 581–581.
- [19] Zhao, W., Chellappa, R., Phillips, P. J., and Rosenfeld, A. Face recognition: a literature survey. *Acm Comput. Surv.* 35, 4 (dec 2003), 399–458.
- [20] Zhou, D., Yang, X., Peng, N., and Wang, Y. Improved-LDA based face recognition using both facial global and local information. *Pattern Recogn. Lett.* 27, 6 (Apr 2006), 536–543

6. APPENDIX

Table 3.1: Recognition rate vs. number of training samples per person (ORL)

No. of training Samples per person	2	3	4	5	6	7	8	Average
EigenFace[16]	82	84	85	86	87	88	90	86
FishersFace[1]	76	83.5	86.5	89	92.5	94	95	88.07
Recognition accuracy (%) (Proposed)	92.25	93.16	94.75	96.5	98.58	98.64	99.68	96.22

Table 3.2: Recognition rate vs. number of training samples per person (YALE)

No. of training samples per person	2	3	4	5	6	7	8	9	10	Avg
Recognition accuracy (%) (Proposed)	91.33	93.55	96	98	98.5	98.14	98.34	98.43	98.64	96.77

Table 3.3: False Recognition rate vs. number of training samples per person (ORL)

No. of training samples per person	2	3	4	5	6	7	8	9
DPA, [3]	0.159	0.098	0.065	0.048	0.038	0.027	0.029	0.019
Pm2+DPA, [3]	0.17	0.10	0.082	0.062	0.05	0.045	0.039	0.03
Proposed	0.14	0.09	0.064	0.045	0.033	0.025	0.025	0.012

Table 3.4: False Recognition rate vs. number of training samples per person (YALE)

No. of training samples per person	2	3	4	5	6	7	8	9	10
DPA, [3]	0.296	0.247	0.219	0.193	0.172	0.16	0.141	0.124	0.12
Pm2+DPA, [3]	0.26	0.219	0.19	0.17	0.16	0.148	0.137	0.123	0.12
Proposed	0.25	0.2	0.16	0.16	0.15	0.14	0.12	0.12	0.11