Dual Objective Feature Selection and Scaled Euclidean Classification for Face Recognition

Siddharth Srivatsa
Student, Dept. of
Electronics and Comm.
M.S. Ramaiah Inst. of Tech.
Bangalore - 560054, India
srivatsasiddharth@gmail.com

Prajwal Shanthakumar
Student, Dept. of
Electronics and Comm.
M.S. Ramaiah Inst. of Tech.
Bangalore - 560054, India
praj.s.92@gmail.com

K. Manikantan
Assistant Professor, Dept. of
Electronics and Comm.
M.S. Ramaiah Inst. of Tech.
Bangalore - 560054, India
kmanikantan@msrit.edu

S. Ramachandran
Professor, Dept. of
Electronics and Comm.
S.J.B. Inst. of Tech.
Bangalore - 560060, India
ramachandr@gmail.com

***Abstract*—The statistical description of the face varies drastically with changes in pose, illumination and expression. These variations make face recognition (FR) even more challenging. In this paper, two novel techniques are proposed, viz., *Dual Objective Feature Selection* to learn and select only discriminant features and *Scaled Euclidean Classification* to exploit within-class information for smarter matching. The 1-D discrete cosine transform (DCT) is used for efficient feature extraction. A complete FR system for enhanced recognition performance is presented. Experimental results on three benchmark face databases, namely, Color FERET, CMU PIE and ORL, illustrate the promising performance of the proposed techniques for face recognition.**

***Index Terms*—Face recognition, feature extraction, feature selection, classifier, discrete cosine transform**

I. Introduction

Face recognition (FR) is the task of face identification by a computer system trained using similar images/video frames of the subject's face. FR has a wide range of security, surveillance and commercial applications. An early survey [1] in FR established that statistical (pixel intensity based) approaches outperform purely geometric approaches (measurement of distance between eyes, shape of the chin, etc.). Comprehensive surveys of recent FR methods, broadly classifiable as holistic (usage of statistics of the face as a whole) and local (extraction of features such as eyes, mouth and nose and usage of their local statistics) can be found in Refs. [2][3]. A comparative evaluation of subspace methods for face recognition is presented in Ref. [4].

The two dimensional (2D) discrete cosine transform (DCT) has been used in Ref. [5] for holistic FR. In this paper, we use the rather unconventional 1-D DCT. After transformation, proper feature selection is critical. A small, optimal subset of features that can discriminate between subjects is crucial. A large feature subset with redundancies and features whose within-class variation is significant compared to inter-class separation, results in compromised performance and greater computational cost.

With this background, the proposed contributions of this paper for enhanced face recognition are:

- *Dual Objective Feature Selector* to learn and select features that maximize inter-class separation and minimize within-class variation.
- *Scaled Euclidean Classifier* to exploit information of within-class variation of the selected features for smarter classification.

The FR system including pre-processing methods, 1-D DCT feature extraction, proposed feature selection and proposed classification is detailed in Section II, the experiments and results are discussed in Section III and conclusions are drawn in Section IV.

II. Discussion of the proposed FR system

The process flow in the proposed FR system is shown in Fig. 1. The system is first trained with images of the subjects and a feature vector for every image is stored in the gallery. The test image to be matched to one of the subjects is subjected to the same processes and a classifier is used for subject identification.

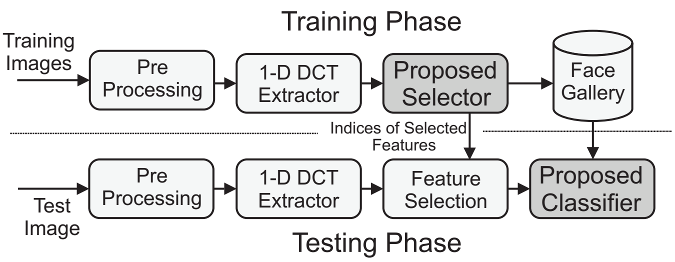

Fig. 1. Process flow in the proposed FR system

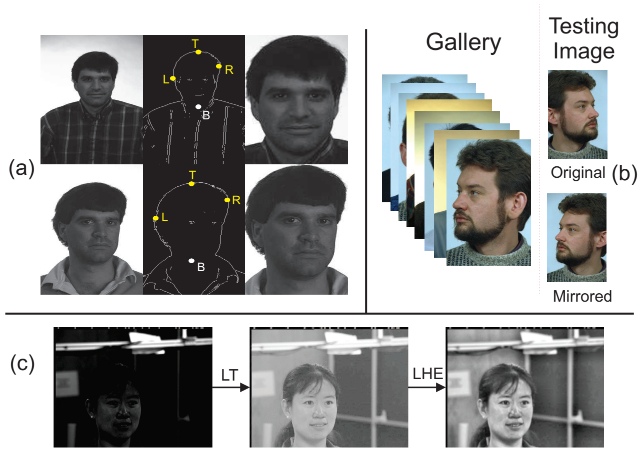

Fig. 2. The pre-processing methods. (a) illustrates scale norm. for 2 images from Color FERET, (b) shows that mirrored test image has a better chance of match with the subject image in the gallery, (c) demonstrates illumination neutralization for an image from CMU PIE

Images are pre-processed to combat variations in illumination, pose, expression and scale (occlusion, background clutter, etc. are not tackled in this paper). The methods used are:

- Scale normalization using edge detection to neutralize the distance from which the face image has been captured. The edges in the facial image are first detected. The image is scanned and the leftmost (L), rightmost (R) and topmost (T) pixels are obtained. The bottommost (B) pixel is computed using knowledge of the aspect ratio of the face. The region between the four points is extracted as the face. The process is shown for two facial images of different scale in Fig. 2a.
- Mirroring of the test image and using it in addition to the original to make recognition more robust to pose variations. Consider, for instance, that the training image of a subject has a right profile. If the test image of the same subject is left profile, it stands a poor chance of matching. Mirroring generates the exact pan with the opposite profile increasing the possibility of a positive match. This is illustrated in Fig. 2b.
- Log transform (LT) + local histogram equalization (LHE) [6] to neutralize illumination (Fig 2c). In facial images, using LT drastically increases pixel intensities in the dark regions and brings out details while the bright regions approximately retain their pixel intensity. This is followed by LHE to increase contrast.

Once pre-processing is completed, the image is raster scanned (row-wise concatenated) to form a large row vector. The 1-D DCT is performed on this row vector to convert the spatial image into the frequency domain. The image is now de-correlated and represented in terms of amplitudes of DCT coefficients. The first 2,500 coefficients are preserved while the others are discarded. (This was experimentally found to be sufficient.) This is done to reduce the computational burden on the selector. From the extracted coefficients, selection of optimal features is performed by the proposed feature selector.

A. Dual Objective Feature Selector

The primary requirement is to define a criterion to evaluate the *fitness* (quality) of features. The logical choice is to select features that vary significantly between subjects. However, a feature can truly distinguish between subjects only if it's variation among images of the same subject is insignificant compared to the variation between subjects, i.e., features with large *inter-class standard deviation (ICSD)* and small *within-class standard deviation (WCSD)*. For the k^{th} feature,

$$fitness_k = \frac{ICSD_k}{WCSD_k} \tag{1}$$

$$ICSD_k = \sqrt{\frac{1}{S}\sum_{i=1}^{S}(\mathrm{M}_{k_i} - \mathrm{M}_k)^2} \tag{2}$$

$$WCSD_k = \sum_{i=1}^{S}\frac{1}{S}\sqrt{\frac{1}{T}\sum_{j=1}^{T}(f_{k_{ij}} - \mathrm{M}_{k_i})^2} \tag{3}$$

S is the no. of subjects,
T is the no. of training images per subject,
$f_{k_{ij}}$ is the amplitude of the k^{th} DCT coefficient in the j^{th} image of the i^{th} subject,
M_{k_i} (class mean) is the k^{th} feature's mean for the i^{th} subject,

$$\mathrm{M}_{k_i} = \frac{1}{T}\sum_{j=1}^{T} f_{k_{ij}} \tag{4}$$

M_k (grand mean) is the k^{th} feature's mean for the database,

$$\mathrm{M}_k = \frac{1}{S}\sum_{i=1}^{S}\mathrm{M}_{k_i} \tag{5}$$

The fitness is computed for all 2,500 extracted features. The selector operates in two steps to select discriminating features. In Step 1, the region of existence of such features is established. In Step 2, individual features are selected from the region identified in Step 1.

1) Identifying the region of high fitness features: Crudely, the aim is to establish whether high quality features are concentrated in the lower coefficients or spread out into the higher end of the spectrum as well. Specifically, the key is to identify the *last* DCT coefficient satisfying a suitably defined fitness criterion and selecting the region of lower frequency features up to that point for Step 2. The criterion is empirically found to be 70 % of the highest fitness in the distribution. In Fig. 3a, the *last* feature satisfying the requirement is coefficient 193, in Fig. 3b, it is coefficient 2319 and in Fig. 3c, it is coefficient 225.

2) Selection of individual features: The task is to pick out the individual features from the region selected in Step 1. While a high threshold was used in Step 1 to establish the coefficient up to which features are to be considered, a lower threshold is employed in Step 2. This is because classifiers require a sufficiently large number of features for reliable classification and a high threshold results in very few features. Any individual feature in the region selected after Step 1 satisfying the new threshold is selected. The new *threshold* (shown in Figs. 3d, 3e and 3f), which is a function of the width of the region selected in Step 1 and the statistics of the distribution is empirically found to be,

$$threshold = Mean + \left(Std.Deviation \times \frac{last}{1000}\right) \tag{6}$$

Also, the *threshold* is limited as,

$$threshold_{\max} = Mean + Std.Deviation \tag{7}$$

There is one other point of consideration. When the width of the region selected in Step 1 is large (Fig. 3b), adjacent features represent similar information. A constraint on the *separation* between selected features can be imposed to restrict the number of features selected without sacrificing performance (for instance, all features satisfying the *threshold* in Fig. 3e won't be selected). In the proposed system, if *last* is greater than 1,250, a minimum *separation* of 1 is used while there is no constraint if *last* is less than 1,250.

B. Scaled Euclidean Classifier

Nearest neighbor classifiers choose as the match, the subject whose feature vector in the gallery is least distant from the feature vector of the test image. In the proposed system, Euclidean distance is used as the metric.

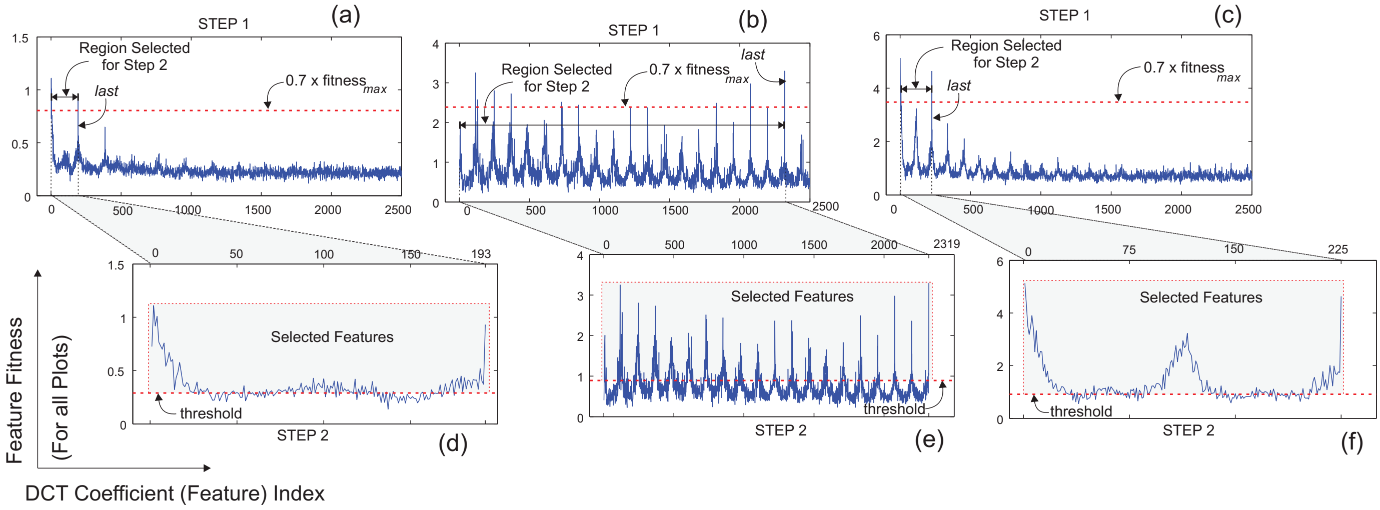

Fig. 3. Feature selection; (a) and (d) illustrate the two step process for Color FERET, (b) and (e) for CMU PIE, (c) and (f) for ORL

If f_k is the k^{th} feature from the test image, g_k the k^{th} feature of a vector in the gallery and F the number of selected features,

$$Euclidean\ Dist. = \sqrt{\frac{1}{F}\sum_{k=1}^{F}(f_k - g_k)^2} \tag{8}$$

We propose scaling the classifier to exploit knowledge of within-class variation. For every feature, the distance between the test vector and a subject vector in the gallery is calculated with respect to the average within-class separation of the feature. In essence, the features are scaled using the average within-class information and this is formulated in Eq. 9. This increases separation between two most likely subjects as illustrated in Fig. 4.

$$Scaled\ Euclidean\ Dist. = \sqrt{\frac{1}{F}\sum_{k=1}^{F}\frac{(f_k - g_k)^2}{WCSD_k}} \tag{9}$$

III. Discussion of Experiments and Results

Experiments were conducted to evaluate the system against pose and scale variance (Color FERET), illumination variance (CMU PIE) and expression variance (ORL). Custom databases were created to set up challenging conditions. The system was realized on MATLAB [10] and tests run on a PC powered by the Intel Core i7 Processor clocking 2.3 GHz with 8 GB of RAM. The extractor, selector and classifier used on all databases is the same while the preprocessing methods used on each database is outlined in Table I.

A. Description of databases and results

Color FERET [11] contains 11,338 color facial images of 994 individuals. A custom database was created by choosing 20 images each from 35 randomly selected subjects, with the image size being 384×256. The variation in scale and pose for one of the subjects is shown in Fig. 5a. The background is plain and light.

CMU PIE [12] consists of 41,368 color facial images of 68 people, each under 13 poses, 43 illumination conditions, and 4 expressions. The image size is 384×256. A custom database was created with 10 images each (Fig. 5b) for 20 subjects, with extreme lighting conditions and varying expressions but no pose variation.

ORL [13] consists of 10 gray-scale facial images, (Fig. 5c) each of 40 distinct subjects. The size of each image is 92×112. Images are upright and frontal with tolerance for some side movement and lighting variation. The background is plain and dark. The variation is in facial expressions and facial details.

Recognition rate (RR = positive hits/total test images), dimensionality (no. of features), training time and testing time/image were the metrics for evaluation. The training and

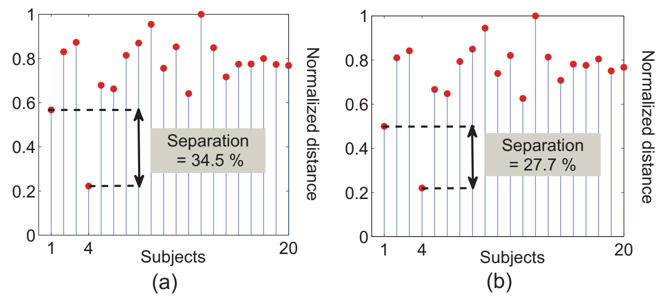

Fig. 4. Using the proposed Scaled Euclidean Classifier (a) results in a larger separation between the correct match and the nearest false subject compared to using the conventional euclidean classifier (b)

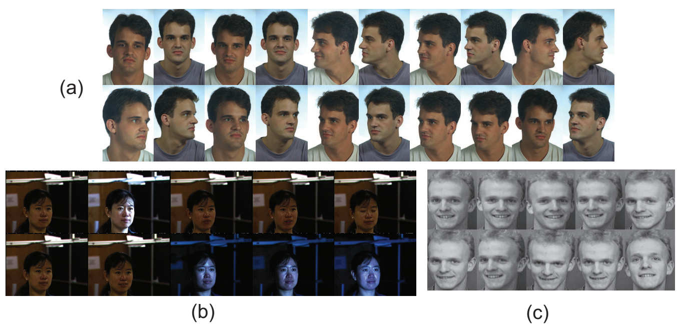

Fig. 5. Sample images. (a) Color FERET, (b) CMU PIE, (c) ORL

TABLE I
PRE-PROCESSING FOR COLOR FERET, CMU PIE AND ORL

Database	Scale Norm.	Mirror	LT + LHE
Color FERET	✓	✓	
CMU PIE			✓
ORL		✓	

TABLE II
EXPERIMENTAL RESULTS ON COLOR FERET, CMU PIE AND ORL

Database	Tr. : Te.	Avg. RR (%)	Features Selected	Training Time (s)	Testing Time / Image (ms)
Color FERET	4:16	83.70	169	2.20	17.80
	8:12	92.25	154	4.41	19.60
	12:8	95.71	162	6.56	21.40
	16:4	97.86	166	8.67	23.00
CMU PIE	2:8	92.75	148	1.19	26.10
	4:6	99.92	227	2.21	28.00
	6:4	100	220	3.24	28.40
	8:2	100	210	4.33	30.00
ORL	2:8	78.90	128	0.46	7.30
	4:6	95.25	153	0.89	8.40
	6:4	97.81	169	1.33	9.50
	8:2	98.87	175	1.76	10.60

TABLE III
COMPARISON WITH OTHER FR SYSTEMS

Color FERET (Tr. : Te. = 8 : 12)		ORL (Tr. : Te. = 5 : 5)	
System	RR (%)	System	RR (%)
DFT + DCT + BPSO [14]	80.00	DCT Fusion + BPSO [15]	95.40
DWT + Laplacian Gradient Masking [7]	85.71	Probabilistic Decision based Neural Network [8]	96.00
Intensity Mapped Unsharp Masking + LOG Filtering + DWT + BPSO [9]	86.45	DFT + DCT + BPSO [14]	97.10
Proposed method	**92.25**	**Proposed method**	**97.20**

Note: BPSO - Binary Particle Swarm Optimization

testing sets were mutually exclusive. The performance of the system for a given test image depends on the training set. Taking this into account, the recognition rate and the number of features selected were averaged over 20 trials of randomly selected training sets. The results are tabulated for the three databases for various training (Tr.) to testing (Te.) ratios in Table II. Dimensionality reduction is illustrated in Fig. 6a. An example of performance gain in using the proposed classifier is shown in Fig. 6b. The recognition performance improves as training is increased. Obviously, training time also increases. However, there is an increase in testing time/image as well with increasing training images due to greater number of distance computations. Thus, there is an accuracy-computation time trade-off. Nonetheless, it can be seen that the number of features required to represent a face is much smaller than the size of the original image. This means that the space required to store faces in the gallery also decreases by the same factor. In addition, a small number of features means lesser number of Euclidean computations and results in a faster system.

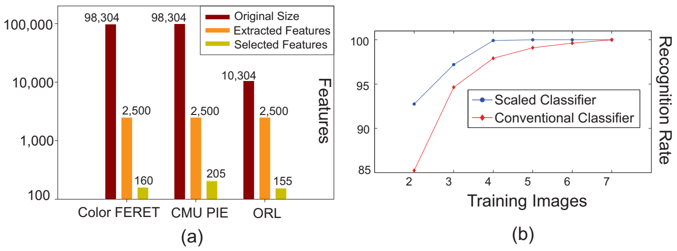

Fig. 6. Dimensionality reduction is graphically illustrated in (a). Performance gain in using the proposed classifier is shown in (b) for CMU PIE database

B. *Comparison with other FR methods*

The system's performance against other FR systems evaluated on similar custom databases is summarized in Table III.

IV. CONCLUSIONS

Unique methods for feature selection and classification, namely, Dual Objective Feature Selection and Scaled Euclidean Classification, have been proposed in this paper and implemented on MATLAB. It can be inferred from Table II that high recognition rates and reduced feature subsets have been obtained by using only discriminant features which contribute towards recognition. The results obtained on applying the proposed method to three face databases that have variations in pose (Color FERET), illumination (CMU PIE) and expression (ORL) illustrate the robustness of the system. Reduction in the number of features used for classification decreases testing times, improving the speed of the FR system. Experiments have been conducted in a constrained environment with the testing images selected from the same database as the training set. Future work shall focus on adapting the system to unconstrained environments. An interesting observation is that discriminant features span a larger spectrum in absence of pose variations while only lower frequency features were found to be discriminating in the presence of pose variations. However, the finding and it's implications require further study.

REFERENCES

[1] R. Brunelli, T. Poggio, *Face Recognition: Features versus Templates*, Pattern Analysis and Machine Intelligence, IEEE Trans. on, vol. 15, pp. 1042-1052, 1993.

[2] W. Zhao, R. Chellappa, P.J. Phillips, A. Rosenfeld, *Face Recognition: A Literature Survey*, ACM Computing Surveys, vol. 35, pp. 399-458, 2003.

[3] X. Zhang, Y. Gao, *Face Recognition across Pose: A Review*, Pattern Recognition, vol. 42, pp. 2876-2896, 2009.

[4] G. Shakhnarovich, B. Moghaddam, *Face Recognition in Subspaces*, Chapter 2, Handbook of Face Recognition, Springer-Verlag London Ltd., pp. 19-49, 2011.

[5] Z.M. Hafed, M.D. Levine, *Face Recognition using the Discrete Cosine Transform*, International Journal of Computer Vision, vol. 43, pp. 167-188, 2001.

[6] R. Szeliski, *Image Processing*, Chapter 3, Computer Vision: Algorithms and Applications, pp. 87-180, 2010.

[7] N.N. Murthy, Raghunandana R. , K. Manikantan, *Face Recognition using DWT Thresholding based Feature Extraction with Laplacian Gradient Masking as a Pre-Processing Technique*, in Proc. of CUBE - International Information Technology Conference, 2012, pp. 82-89.

[8] S.H. Lin, S.Y. Kung, L.J. Lin, *Face Recognition/Detection by Probabilistic Decision based Neural Network*, IEEE Trans. Neural Network, vol. 8, pp. 114-132, 1997.

[9] G.S. Yaji, S. Sarkar, K. Manikantan, S. Ramachandran, *DWT Feature Extraction based Face Recognition using Intensity Mapped Unsharp Masking and Laplacian of Gaussian Filtering with Scalar Multiplier*, in Proc. of 2nd International Conference on Communication, Computing & Security, 2012, pp. 475-484.

[10] MATLAB: www.mathworks.com

[11] Color FERET: http://www.nist.gov/itl/iad/ig/colorferet.cfm

[12] CMU PIE: http://www.ri.cmu.edu/research_project_detail.html?project_id=418&menu_id=261

[13] ORL: http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html

[14] G.M. Deepa, R. Keerthi, N. Meghana, K. Manikantan, *Face Recognition using Spectrum based Feature Extraction*, Applied Soft Computing, vol. 12, pp. 2913-2923, 2012.

[15] Prathik P. , R.A. Nafde, K. Manikantan, S. Ramachandran, *Feature Extraction using DCT Fusion based on Facial Symmetry for enhanced Face Recognition*, in Proc. of ICCICT - International Conference on Communication, Information and Computing Technology, 2012.