Journal of Advanced Computer Science and Technology, 1 (4) (2012) 266-283

www.sciencepubco.com/index.php/JACST

Face Recognition using Block-Based DCT Feature Extraction

K Manikantan[1], Vaishnavi Govindarajan[1], V V S Sasi Kiran[1], S Ramachandran[2]

[1]Department of Electronics and Communication Engineering,
M S Ramaiah Institute of Technology, Bangalore, Karnataka, India 560054
E-mail: kmanikantan@msrit.edu
E-mail: vaish.india@gmail.com
E-mail: sasikiran.f4@gmail.com
[2]Department of Electronics and Communication Engineering,
S J B Institute of Technology, Bangalore, Karnataka, India 560060
E-mail: ramachandr@gmail.com

Abstract

Face recognition (FR) with reduced number of features is challenging and energy based feature extraction is an effective approach to solve this problem. However, existing methods are hard to extract only the required low frequency features, which is important for capturing the intrinsic features of a face image. This paper proposes a novel *Block-Based Discrete Cosine Transform* (BBDCT) for feature extraction wherein each 8×8 DCT block is of adequate size to collect the information within that block without any compromise. Individual stages of FR system are examined and an attempt is made to improve each stage. A Binary Particle Swarm Optimization (BPSO)-based feature selection algorithm is used to search the feature vector space for the optimal feature subset. Experimental results show the promising performance of BBDCT for face recognition on 4 benchmark face databases, namely, ORL, Cropped UMIST, Extended Yale B and Color FERET databases. A significant increase in the overall recognition rate and a substantial reduction in the number of features, are observed.

Keywords: *Discrete Cosine Transform, Face Recognition, Feature Extraction, Feature Selection, Particle Swarm Optimization*

1 Introduction

Face recognition (FR) is a challenging field owing to its complexity and extensive usage in applications in fields such as forensics, vigilance, Law enforcement, user access, human computer interaction and for various other security purposes. It is superior over fingerprints and other biometrics since it works without the involvement and knowledge of the individual concerned [1]. Many commercial systems thus use face recognition. The process of Face Recognition comprises of Face Detection, feature extraction and verification or identification [1, 2]. This paper highlights the extraction and identification stages in the FR process. Many face recognition algorithms have been developed. However due to position, pose, noise and illumination variations in the input image, the success rate of the algorithms is not deterministic [3]. This has resulted in development of innumerable robust techniques such as background removal, illumination normalization and others which support the algorithm to combat the undesirable effects and improve the success rate [4]. An exhaustive survey on FR techniques is given in Ref. [1]. Feature extraction using block based DCT involves dividing the image into blocks of uniform size and isolating the most relevant features of each block [5]. The DC component or low frequency components contain the maximum relevant information useful for face recognition with illumination variation effects, while the high frequency components correspond to finer details like edges and expressions which are vulnerable to pose and expression variations. Thus, local analysis using block based DCT is performed and the necessary features are combined to represent the extracted features. Further, the BPSO structure based on swarm intelligence is used to reduce the feature subset, and improve the performance of the system in terms of computational time reducing the complexity and also the number of selected features.
For enhanced Face Recognition, this paper proposes the following new ideas:

1. Block Based DCT for optimal feature extraction.

2. Optimal features selected based on Energy Pruning levels which are further reduced through BPSO.

This paper is organized as follows: Pre-processing techniques employed and conventional DCT based feature extraction in section 2, the proposed algorithm for optimal feature extraction using block based DCT approach and feature selection using BPSO in sections 3 and 4, and Experimental results in section 6.

2 Preliminary

2.1 Image Pre-processing techniques

The main objective of pre-processing is to make the images more suitable for the application of algorithm in any system. The pre-processing techniques are situation specific and they improve the success rates drastically when used appropriately. The following are the pre-processing techniques applied in our algorithm:

- *Scaling:* It is expected that the images of the probe and the face gallery to be scaled to the same size as the correlation between the images of the same size is maximum [1].Face gallery corresponds to the set of training images and probes are the test images to be identified. The scaling factors for different databases are mentioned in section 5.

- *Histogram Equalization:* It is an illumination Normalization technique. Images which have unequal value of brightness can be improved by using Histogram equalization process.Histogram equalization is a non linear process aimed to highlight brightness in a way particularly suited to human visual analysis [6]. The objective of Histogram equalization is to produce equal value of intensity over the entire range of pixels. It automatically minimizes the contrast in areas which are too light or too dark for an image. It is a transformation that considers the accumulative distribution of the original image to generate a resulting image whose histogram is approximately uniform. It is mathematically given by Eq. 1.

$$h(v) = round\left(\frac{cdf(v) - cdf\min}{(M \times N) - cdf\min} \times (L-1)\right) \tag{1}$$

  M, N is the width and the height of the image respectively, CDF is the Cumulative Distribution Function and L is the number of gray levels used.

- *Edge Detection:* Edges correspond to the local variations in the intensity of the image pixel values. In order to preserve local features, despite the influence of lighting, useful for recognition, LoG [7] is applied after histogram equalization. There are various edge extraction techniques like 'Canny', 'Sobel' etc., out of which *Laplacian of Gaussian (LoG)* is used here for illumination invariant database. The edge techniques work their best in the illumination normalized images.

2.2 Discrete Cosine Transform

DCT is the most widely used transform in the image processing applications for feature extraction. The approach involves taking the transformation of the image as a whole and separating the relevant coefficients. DCT performs energy compaction [8]. The DCT of an image basically consists of three frequency components namely low, middle, high each containing some detail and information in an image. The low frequency generally contains the average intensity of an image which is the most intended in FR systems [9]. Mathematically, the 2D-DCT of an image is given by:

$$F(u,v) = \alpha(u)\alpha(v) \sum_{x=0}^{N-1} \sum_{y=0}^{M-1} \cos\left[\frac{\pi u}{2N}(2x+1)\right] \cos\left[\frac{\pi u}{2M}(2y+1)\right] f(x,y) \quad (2)$$

$$\alpha(u)\alpha(v) = \left\{ \begin{array}{l} \sqrt{\frac{1}{N}} for u, v \neq 0 \\ \sqrt{\frac{2}{N}} for u, v = 0 \end{array} \right\} \quad (3)$$

where f(x, y) is the intensity of the pixel at coordinates (x,y), u varies from 0 to M-1, and v varies from 0 to N-1, where M×N is the size of the image.

3 Proposed Block-Based DCT (BBDCT) for Feature Extraction

DCT is optimal in the sense of decorrelating the data and maximizing the energy packed into the lower order coefficients. It reduces computational complexity in implementation. Our proposed technique is a hybrid of feature selection and reduction algorithms. The success of a face recognition algorithm mainly depends on selecting the relevant features while ignoring the others which tend to vary due to external disturbances and other factors [1]. Block-based DCT (BBDCT) is an efficient approach in feature extraction over the conventional DCT, which operates on the image as a whole.

In the proposed system architecture as shown in Fig. 1, the image is divided into blocks of 8×8. When the dimensions of the image are not integral multiples of 8, the algorithm pads zeroes along the rows and/or columns to get the integral number of blocks. The three reasons for selecting the block-size to be 8×8 are:

- It is an adequate size in order to collect the information from all the regions of the face without a compromise [5].
- Useful image contents vary gracefully within an 8×8 image block.

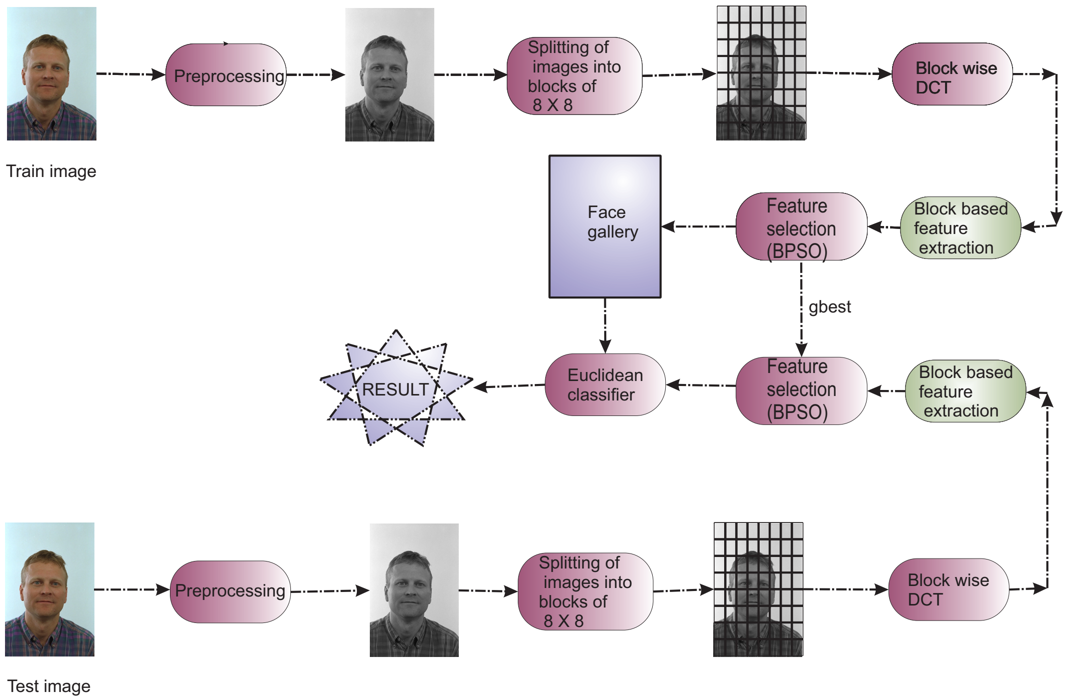


Figure 1: Block diagram of the proposed FR system

- To favor the hardware implementation of the algorithm as this size is best suited for meeting both the processing and timing constraints of most of the real time processors.

2D-DCT is applied to the blocks of 8×8 to result in 64 coefficients. A 2D-DCT is preferred since image is two dimensional, which makes the correlation approach convenient and also it is tedious to compute 1D-DCT coefficients along each dimension of the image and cascade them to form the 2D-DCT matrix. The 64 coefficients thus obtained per block are arranged in a matrix of size 8×8 in raster scan order as explained later.

The first coefficient in this matrix is known as the *DC component, representing the average intensity of an image*, while the rest are the AC coefficients corresponding to high frequency components of the image. It is proved that *High-frequency information by itself is insufficient for good face recognition performance* [10]. The elimination of the high frequency components also causes the image to be robust to scale variations which is required in FR systems.

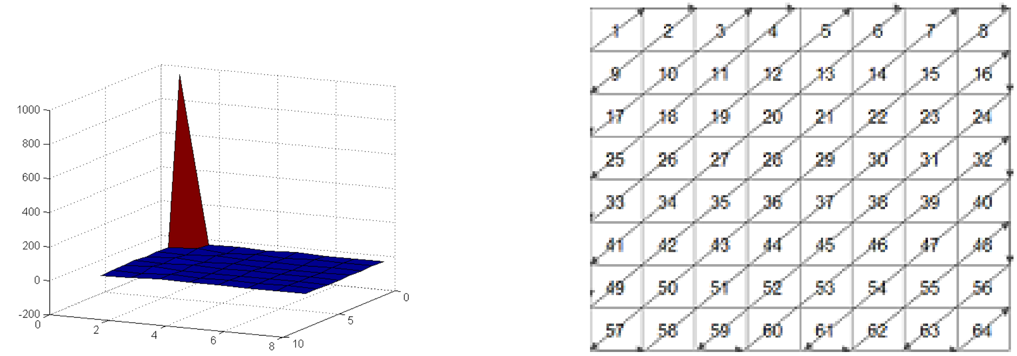
The following theory supports the superiority of Block-Based DCT approach: In an image, the top left entry referred to as the DC coefficient, stores the amplitude and the base frequency. It is practically observed that the DC coefficient carries 95% of the energy. An experimental proof of the same,is given in Table 1 Amplitudes are directly related to the energies which carry the information in an image. Outcome of DCT is the collection of amplitudes

Table 1: Energy content at different pruning levels for each database

Pruning levels	ORL	UMIST	FERET	Extended Yale B
	1.0e+005×	1.0e+004×	1.0e+006×	1.0e+003×
Level-0	9.6826	2.7308	1.2432	3.1220
Level-1	0.0003	0.0002	0.0002	0.0767
Level-2	0.0003	0.0000	0.0001	0.0050
Level-3	0.0002	0.0000	0.0000	0.0022
Level-4	0.0001	0.0003	0.0000	0.0008
Level-5	0.0000	0.0003	0.0000	0.0013
Level-6	0.0002	0.0000	0.0000	0.0010
Level-7	0.0000	0.0000	0.0000	0.0000

at the more significant lower frequencies (top left quadrant) and lesser entries at higher frequencies [11]. Thus, it is essentially the low frequency components, which play a key role in FR systems [12]. This observation is the basis for the development of our algorithm. Thus, instead of taking the correlated low frequencies at the top left corner of the image as a whole, it is useful to pick up the low frequencies at each subsection of the image so that they are more de-clustered and uncorrelated to highlight the important features of each subsection of the image. We refer to this subsection as a block owing to its 8×8 size and appearance. In this paper, the above technique is proved to be successful in the field of face recognition. The smoother variations in an image (or the low frequencies) are present at the top left corner of the image. Though these are the best for the purpose of recognition, these components are susceptible to illumination and environmental variations as the perception of smoothness is not uniform under different conditions [2].

Edges or the high frequencies on the other hand are abrupt across the surface of the image, and are the low energy content components of the image. It is because of this fact that the recognition of faces under non-uniform conditions should also include a minimal amount of edges. One way to handle this problem is to make the block size smaller when the details or transitions in an image are more, so that across that block in the entire image, the transitions are smooth. Smaller the block size, more the correlation between adjacent blocks since correlation in general decreases with distance in an image. In an image with lesser details, too small size of the blocks will result in too many blocks, with redundant features, which does not serve the basic purpose of decorrelating the data [13]. Selecting blocks with larger block size would decorrelate the features but reduce the overall number of important features required for recognizing a face. Thus, selection of an optimum block size is essential. In an image with pixel values in the range of [0,255] (grayscale image), for a block size of 8×8, the DCT values are in the range of $\pm$ 255×8×8.



(a) Magnitude of DCT of a single block in the image

(b) Raster scan order

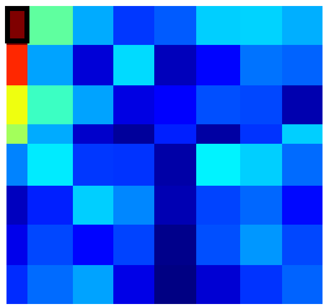
Figure 2: 8×8 operations

Within each DCT block of the image, the pixels differ from a low variation to a high variation in a zig-zag pattern, better known as the raster scan order, which is the basis for selecting the number of features in the selected blocks. A typical DCT magnitude plot is shown in Fig. 2a and the corresponding DCT coefficients' order is shown in the Fig. 2b.

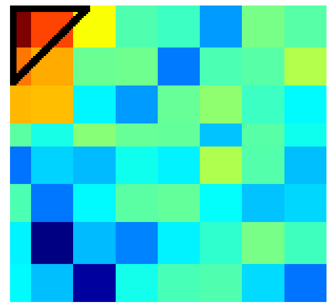
The raster scan order which is the basis for selection of number of coefficients within the block size of 8×8 as mentioned earlier consists of arrangement of DCT coefficients as a matrix is as shown in Fig. 2b. The primary diagonal elements in the matrix are the pruning levels which indicate the order of importance of the DCT coefficients based on the energy carried by that coefficient. The pruning level is chosen so that the energy carried by the coefficient is significant enough for computation. Experiments which are performed using our algorithm involves inclusion of coefficients of various pruning levels-level 0, level 1, and level 2 and the variation in the performance of the system with this is observed. Pruning levels depend on the nature of the images owing to the fact that the distribution of energy in each image varies according to the conditions to which is subjected. Thus, the results obtained varied for databases as described in Section 4. Various Pruning levels (upto Level-2) are as shown in Fig. 3.

4 Binary Particle Swarm Optimization (BPSO) based Feature Selection

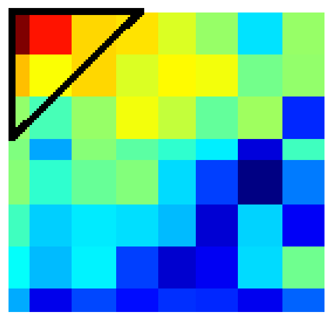
The extracted features are further reduced using swarm intelligence based BPSO technique.



(a) Level-0



(b) Level-1



(c) Level-2

Figure 3: Pruning Levels

The particle Swarm algorithm introduced by Eberhart and Kennedy (1995) is used as an optimization technique in real number space [14]. It uses the approach of dividing the entire problem into a set of particles where every particle represent a possible solution. The position and velocity of every particle is updated during every iteration based on the current particles best position and velocity and also the position and velocity of the best in the swarm. The updating is done based on the following equations Eq. 4 and Eq. 5 [15]:

$$V_i^{t+1} = \omega \times V_i^t + c_1 \times rand_1 \times (p_{i\_best} - X_i^t) + c_2 \times rand_2 \times (g_{best} - X_i^t) \quad (4)$$

$$X_i^{t+1} = X_i^t + V_i^{t+1} \quad (5)$$

where i is the particle index, varying from 1 to N (Number of particles), ω represents the inertia, c_1 and c_2 are the cognitive parameters, g_{best} is the global best position and p_{best} is the particle best position. Also rand() is a uniform random number in the range 0-1.

X_i^{t+1} is the predicted position, X_i^t is the current position and V_i^{t+1} is the estimate of velocity calculated from the previous equation.

The binary variant of PSO given in Eq. 6 has the particle velocity function used as the probability function for the position update[16]. If

$$rand_3 < \frac{1}{1 + e^{-V_i^{t+1}}} \tag{6}$$

then X_i^{t+1} =1, else X_i^{t+1} =0. A value of 1 for the position indicates the selection of the particular feature while a 0 indicates rejection from the required feature set.

For our experimentation, we have chosen c_1=2, c_2=2 and ω=0.6.

Each particle representing a candidate solution is evaluated based on the fitness function, which is a measure of class separation calculated using the Eq. 7[14].

$$F = \sqrt{\sum_{i=1}^{L} (M_i - M_0)^t (M_i - M_0)} \tag{7}$$

where M_i is the mean of the corresponding classes and M_o is the grand mean. Here, classes correspond to the different subjects (L).

It is shown in the succeeding sections that the usage of Particle Swarm Optimization algorithm along with block based DCT technique reduces the feature vector size upto 75%, without affecting the recognition rate.

5 Euclidean classifier

Euclidean classifier is used to find the best match between the train and test images. It is calculated as shown in Eq. 8.

$$D = \sqrt{\sum_{i=1}^{N} (p_i - q_i)^2} \tag{8}$$

p_i and q_i are the coordinates of p and q in the N dimensional space, corresponding to the train and test images. Minimum distance thus corresponds to maximum correlation.

6 Experimental Results

The BBDCT algorithm with BPSO is tested on ORL, Extended Yale B, Color FERET and Cropped UMIST databases. The sample images are as shown in Fig. 4 and the results are averaged for ten iterations. The training and testing images taken randomly over every iteration.

(a) Images from ORL database

(b) Images from subset-5 of extended Yale B database

(c) Images from UMIST database

(d) Images from Color FERET database

Figure 4: Characteristics of each database shown through sample images

The following are the two main test performance criteria in every database taken for experimentation[2]:

- Recognition Rate (RR): It is defined as the number of images identified out of a set of images given to the system for test.
- Features selected: A reduced number of important features selected increase the performance of the system and also reduces the testing time due to the lesser comparisons between the train and the test images.

6.1 ORL database

ORL database [17] consists of 40 distinct individuals with 10 images of each individual of size 112×92 all upright frontal with uniform background. Division into 8×8 blocks, thus, yields 168 blocks.The DC DCT coefficient is taken from each block as mentioned earlier and recognition rate is plotted as a function of the training to testing image ratio. This experiment is also performed with level-2 pruning and the recognition rate is found to be almost the same as that of Level-0 pruning. Since the number of features increase (also the testing time) with the pruning level without affecting the recognition rate, Level-0 (DC coefficient) is superior in terms of performance criteria which proves our theory.

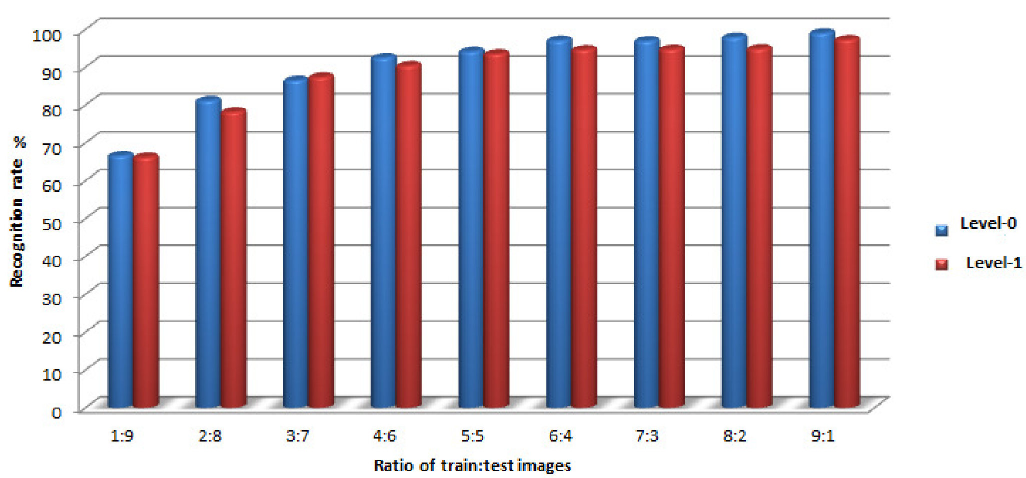


Figure 5: Variation of recognition rate with pruning levels and train:test ratio in ORL database

6.2 Color FERET database

Color FERET database [18] consists of 11338 images in the "smaller" category, with 13 varying poses as shown in the table, with 11 to 50 pose-varied images

Table 2: Color FERET Subsets

Notation	Orientation
fa	regular frontal image
fb	alternative frontal image
pl	profile left
hl	half left, head turned about 67.5 degree left
pr	profile right
hr	half right, head turned about 67.5 degree right
qr	quarter right, head turned about 22.5 degree right
ql	quarter right, head turned about 22.5 degree left
ra	random image, head turned about 45 degree left
rb	random image, head turned about 15 degree left
rc	random image, head turned about 15 degree right
rd	random image, head turned about 75 degree right

per subject. It is characterized by varying poses, facial expressions and facial details, as shown in Table 2. Frontal images (fa,fb)were taken from 40 subjects with 2 images per subject of size 256×384 in our experiments .The images were converted into grayscale and scaled down by a factor of 4 resulting in images of size 64×96. Log transform is applied to improve the contrast of the images. The image is then divided into 96 blocks of size 8×8 each. The DC DCT coefficient is taken from each block as mentioned earlier and recognition rate obtained is 88.9% with a decrease in features upto 75% as compared to conventional DCT.

6.3 Extended Yale B database

Yale B database [19] consists of 28 subjects with 9 images per subject, of size 640×480 with 64 illumination levels, divided into 5 subsets. These are charac-

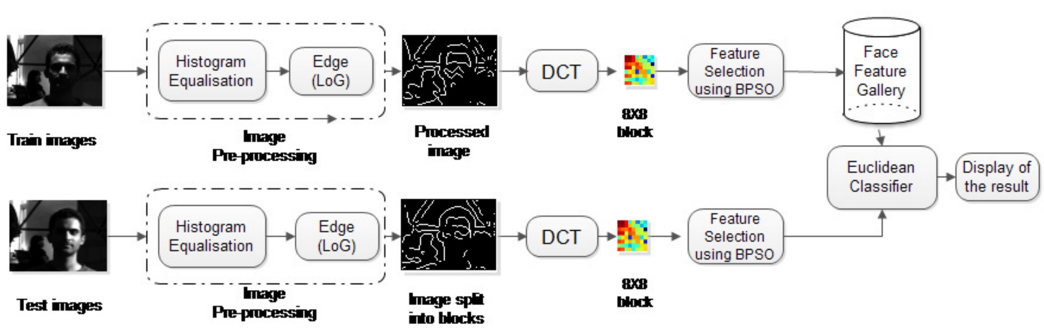


Figure 6: Block diagram of the proposed FR system for Yale B database

Table 3: Extended Yale B Subsets

Subsets	Lighting angle (in degrees)	Number of images for 28 subjects
1	0-12	$7 \times 28 = 196$
2	13-25	$12 \times 28 = 336$
3	26-50	$12 \times 28 = 336$
4	51-77	$14 \times 28 = 392$
5	>77	$19 \times 28 = 532$

terized by uneven background and pose and illumination variations. Subsets are classified based on the lighting angle as shown in Table 3. Subset 5 is considered in our experiments as it happens to be the most challenging of the five. It consists of 28 subjects with 19 images per subject. These images were scaled by a factor of 8 to 80×60 and used with illumination pre-processing Histogram Equalization along with edge detection(LoG). Block division,thus, results in 80 8×8 blocks as shown in Fig. 6.

The DCT coefficients upto Level-2 are taken from each of these blocks and recognition rate is plotted as a function of the training to testing image ratio. Here, it is found that the concentration of energy spans to two pruning levels due to which the Recognition Rate by considering these coefficients is maximum. The results are shown in Fig: 7.

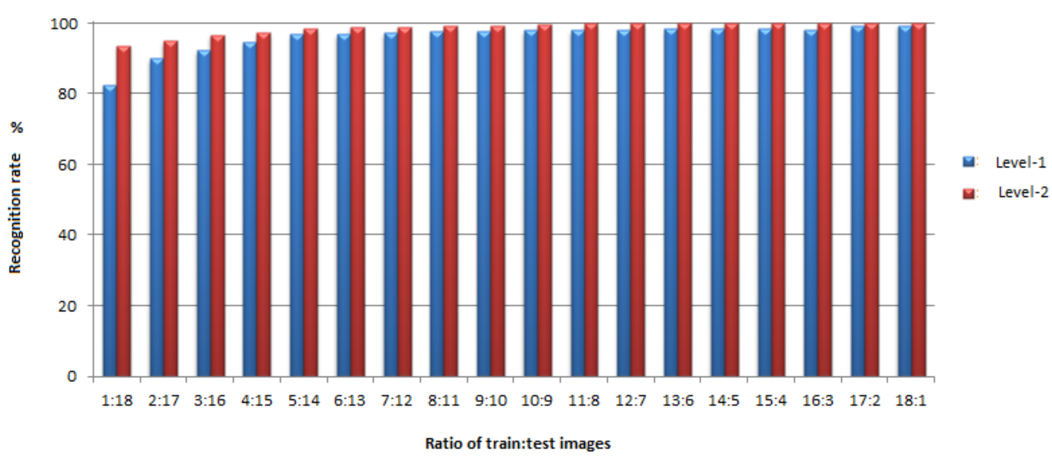


Figure 7: Variation of recognition rate with pruning levels and train:test ratio in extended Yale B database

Table 4: Recognition rates at different pruning levels for Yale B, UMIST and ORL databases

Train:Test	Level	Yale B	UMIST	Train:test	Yale B	UMIST	Train:Test	ORL
1:18	a	82.74	59.61	10:9	99.96	96.11	1:9	66.61
	b	93.15	58.28		99.84	96.22		66.19
2:17	a	90.75	71.44	11:8	99.95	97.56	2:8	81.16
	b	95.15	72.50		99.91	97.00		78.16
3:16	a	94.17	81.22	12:7	99.64	97.07	3:7	86.50
	b	98.75	79.28		100	98.50		87.35
4:15	a	97.64	87.47	13:6	99.75	97.92	4:6	92.50
	b	99.05	85.27		100	98.67		90.40
5:14	a	99.48	87.68	14:5	99.5	98.60	5:5	94.15
	b	99.56	88.75		100	99.12		93.50
6:13	a	99.34	90.09	15:4	99.64	99.37	6:4	97.06
	b	99.86	91.89		100	98.07		94.62
7:12	a	99.38	95.33	16:3	99.76	98.33	7:3	96.92
	b	99.91	93.50		100	99.00		94.75
8:11	a	99.55	96.51	17:2	96.50	99.25	8:2	97.87
	b	99.68	97.28		100	98.25		94.87
9:10	a	99.39	96.09	18:1	99.29	98.50	9:1	99.00
	b	99.86	95.60		100	99.50		97.25

where
a is Level-0 for UMIST,ORL and Level-1 for Yale B
b is Level-1 for UMIST,ORL and Level-2 for Yale B

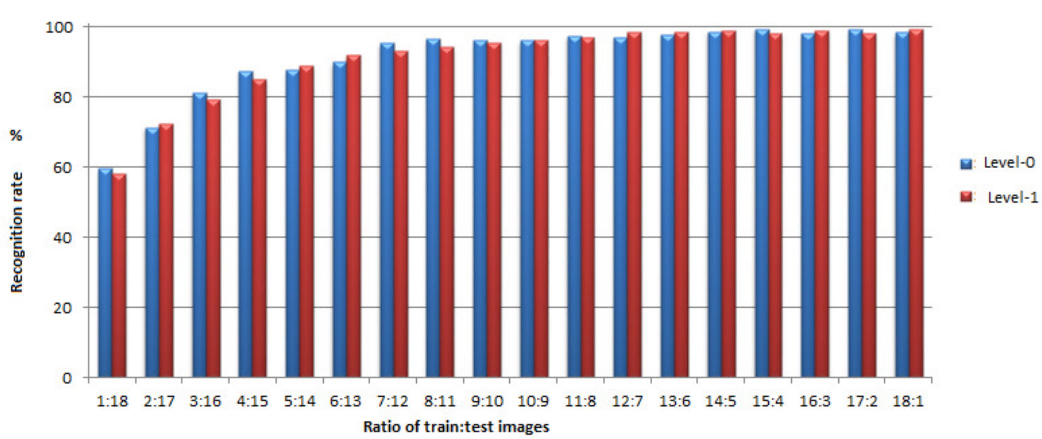


Figure 8: Variation of recognition rate with pruning levels and train:test ratio in cropped UMIST database

Table 5: Comparison with existing FR systems for ORL database (9:1)

Methods	Average recognition rate
ICA [21]	93.80%
Gradient direction [22]	95.75%
Correlation filters [23]	96.25%
Eigen faces [21]	97.50%
Kernel Eigen faces [21]	98.00%
2DPCA [24]	98.30%
Fisher faces [21]	98.50%
Combination [25]	98.50%
Proposed method	99.00%

6.4 Cropped UMIST database

The above algorithm is also tested on Cropped UMIST database [20] having 20 subjects and 19 images per subject, all grayscale with occlusion variation, with no variation in illumination. The size of each image is 112×92, which results in 168, 8×8 blocks for which DCT is applied.

The results obtained by varying the ratio of Training and Testing images for all databases are as shown in the Table 4.

The graphical representation of the results obtained and the number of features selected for all the databases is as shown in Fig. 8.

6.5 Comparison with other FR systems

The number of Training and Testing images are varied in different ratios and the corresponding average recognition rates have been tabulated (Table 4). To compare with other FR systems for ORL database the following experiment is conducted. 9 random images of each person are used for training and the remaining 1 image for testing (9:1). This is repeated for 10 trials and the average recognition rate obtained is compared with other existing techniques in Table 5.

For Extended Yale B database (subset 5, 3:16 ratio), the proposed FR system gives an RR of 98.75% with 110 features as compared to 93.13% and 653 features of the DFT technique proposed in Ref. [26].

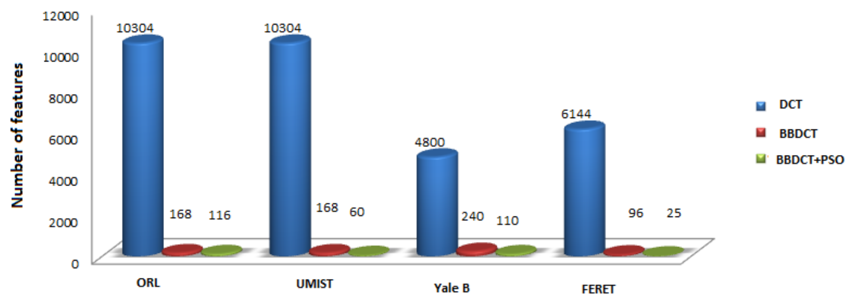


Figure 9: Comparison of number of features in each database using three methods

7 Conclusions

A new approach for flexible FR system is proposed which uses Block Based Discrete Cosine Transform (BBDCT) for feature extraction and a BPSO based feature selection. BBDCT has played a key role in efficient feature extraction which has been the main contributor for the reduced number of features being obtained. A successful attempt has been made to equally handle all image variations (Illumination, facial expressions and pose). The experimental results indicate that the proposed method has performed well under severe illumination conditions with top recognition rates having reached 100% for subset 5 of Extended Yale B. It is also successful in tackling the most challenging task of pose variance in FR with top recognition rates of 99.75% and 93% for UMIST (18:1) and FERET (1:1) respectively. On a PC with Intel i3 2.53 GHz CPU and 3.00 GB RAM, BBDCT costs an average testing time of 100 ms per image of UMIST database when implemented using MATLAB [27]. This may still be a limitation of BBDCT for real time application.

BBDCT and population based optimization algorithm such as BPSO are well suited for execution in parallel stages. This allows the algorithms to be implemented directly *in hardware* and achieve much faster execution times than possible with software.

References

[1] Chellappa *Et.,al*, Face Recognition- Literature Survey ,ACM Computing Surveys,Volume 35 (2003) 399-458.

[2] Mathew Turk, A Randomn Walk through Eigenspace, IEICE Transactions on Information & Systems, Vol. E84-D,No.12 (2001) 1586-1595.

[3] Ramanathan N, Chellappa R, Roy Chowdhury A.K, Facial Similarity Across Age, Disguise, Illumination and Pose, Proceedings of the International Conference on Image Processing (2004) 1999-2002.

[4] Xuan Zou, J. Kittler, K. Messer, Illumination Invariant Face Recognition : A Survey, First IEEE International Conference on Biometrics: Theory, Applications, and Systems (2007) 1-8.

[5] Vikas Maheshkar, Sushila Kamble, Suneeta Agarwal, and Vinay Kumar Srivastava, DCT-Based Reduced Face For Face Recognition, International Journal of Information Technology and Knowledge Management, Vol. 5 (2012) 97-100.

[6] Feature Extraction & Image Processing, Mark S. Nixon, Alberto S. Aguado, Second edition, Elsevier Ltd, 2008.

[7] John Canny, A Computational Approach to Edge Detection, IEEE Transactions on pattern analysis and machine intelligence, Vol. PAMI-8 (1986).

[8] N. Ahmed, T. Natarajan, K.R. Rao, Discrete Cosine Transform, IEEE Transactions on Computers ,Vol. C-23 (1974) 9093.

[9] Digital Image Processing, Rafael Gonzalez, Richard Woods, Edition 3, Pearson Publications, 2008.

[10] Pawan Sinha,Benjamin Balas, Yuri Ostrovsky, Richard Russell,Face Recognition by Humans: Nineteen Results All Computer Vision Researchers Should Know About , In proceedings of IEEE, Vol. 945 (2006).

[11] Z.M. Hafed and M.D. Levine, Face Recognition Using the Discrete Cosine Transform, International Journal of Computer Vision, Vol. 43(3) (2001) 167-188.

[12] F.M. Matos, L.V. Batista, and J. Poel, Face Recognition Using DCT Coefficients Selection, In proceedings of the ACM Symposium on Applied Computing (2008) 1753-1757.

[13] Discrete cosine transform: algorithms, advantages, applications, K.R.Rao.

[14] J. Kennedy, R. Eberhart, Particle Swarm Optimization, In proceedings of the IEEE International Conference on Neural Networks (1995).

[15] Rabab M. Ramadan and Rehab F. Abdel - Kader, Face Recognition Using Particle Swarm Optimization-Based Selected Features,International Journal of Signal Processing, Image Processing and Pattern Recognition Vol. 2, No. 2 (2009).

[16] J.Kennedy and R.C.Eberhart, A Discrete Binary version of the Particle Swarm Optimization, In proceedings of IEEE,International Conference on Systems,Man,Cybernetics,vol. 5 (1997) 4104-4108.

[17] ORL Database: AT&T Laboratories, Cambridge, UK, The ORL Database of Faces (now AT&T The Database of Faces). Available Online: `http://www.cl.cam.ac.uk/Research/DTG/attarchive/pub/data/att_faces.zip`

[18] Color FERET Database: `http://www.nist.gov/humanid/colorferet/home/html`

[19] Athinodoros Georghiades, Peter Belhumeur, David Kriegman, From Few to Many: Illumination Cone Models for Face Recognition under Variable Lighting and Pose, PAMI (2001). Available Online :`http://vision.ucsd.edu/~leekc/ExtYaleDatabase/ExtYaleB.html`.

[20] UMIST Database: The University Of Sheffield, UK, "The UMIST Face Database"(now known as "Sheffield Face Database").Available online: `http://www.sheffield.ac.uk/eee/research/iel/research/face`

[21] M. Yang, Kernel EigenFaces vs. Kernel Fisherfaces: Face Recognition Using Kernel Methods, In proceedings of the Fifth IEEE International Conference on Automatic Face Gesture Recognition (2002) 215-220.

[22] H. F. Chen, P. N. Belhumeur, D. W. Jacobs, In search of illumination invariance, In proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Vol. 1 (2000) 254-261.

[23] B.V.K.V. Kumar, M. Savvides, C. Xie, Correlation pattern recognition for face recognition, In proceedings of the IEEE transactions 94 (2006) 1963-1976.

[24] Jian Yang, David Zhang, Senior Member, IEEE, Alejandro F.Frangi and Jing-yu Yang, Two-Dimensional PCA: A New Approach to Appearance-Based Face Representation and Recognition, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 26, No. 1 (2004).

[25] Raghuraman Gopalan and David Jacobs, Comparing and combining lighting insensitive approaches for face recognition, Computer Vision and Image Understanding 114 (2010) 135-145.

[26] G M Deepa, R Keerthi, N Meghana and K Manikantan, Face Recognition using spectrum based feature extraction, Applied Soft Computing 12 (2012) 2913-2923.

[27] MATLAB: `http://www.mathworks.com`