

Face Recognition using Two Dimension Fractional Discrete Cosine Domain and BPNN

Kumud Arora
Inderprastha Engg. College
Ghaziabad, India.

V.P. Vishwakarma, Ph.D.
GGSIIP University
New Delhi, India

Poonam Garg, Ph.D.
IMT, Raj Nagar
Ghaziabad

ABSTRACT

Face database contains images taken at various instants of the same person. The matching accuracy of spatial features drops significantly both in the presence of noise as well as when the variations in the different instances are high. To attenuate the image variations up to certain extent, image data is transformed from spatial domain to transformed domain. In this paper an effort is made to explore the effect of using fractional order spectrum obtained by the application of 2D FRDCT on the accuracy of face recognition. PCA is used as dimension reduction approach for reducing transformed feature set dimensionality. Reduced feature set is then classified by back propagation neural network classifier. Through the experiments performed on AT&T database it is shown that proposed FRDCT feature set approach gives a recognition accuracy of 94% with BPNN. Comparison is conducted for fractional order feature classification accuracy of AT&T public database with nearest neighbour classification approach. Experimental result shows marked reduction in classification error rate with neural network classification.

General Terms

Face recognition, Fractional DCT.

Keywords

DCT (Discrete Cosine Transformation), BPNN (Back propagation neural network), FRDCT (Fractional Discrete Cosine Transformation), PCA (Principal Component Analysis).

1. INTRODUCTION

Image frequency domain transformation particularly discrete cosine transformation forms part of several standards especially in image compression (MPEG, JPEG, H.23). Discrete cosine transformation (DCT) represents finite sequence of input data points in terms of a sum of cosine functions at different frequencies [13]. Out of four standards of DCT variant, in particular the DCT-II is often used in signal and image processing due to symmetry of its basis matrix [7]. The application of fractional version of discrete cosine transformation came up just few years back [3] whereas the root of this transformation (Discrete Fraction Fourier Transformation) can be traced up to five decades back [11].

Motivated by the encouraging results of Fractional Fourier Transformations (FRFT) in signal processing, Pei and Cariolaro in their work introduced Fractional Discrete Cosine Transform [2-3]. The Fractional discrete cosine transformation like another fractional transformations (Fractional Fourier transformation (FRFT) [10], Fractional Sine Transformation [3], Fractional Wavelet Transformation [12]), integrate the information of time domain and frequency domain at the same time. Pei and Yeh [4] in their work

classified various definitions of fractional cosine transformation according to the methodologies used for calculations. The most referred FRDCT definition in various research papers [19-21] is defined in [2]. This paper uses the Eigen decomposition of the Discrete Cosine Transform (DCT) kernel. Fractional DCT is generalized form of conventional DCT with an additional free parameter and using this free parameter, DFRDCT paves its way into all the applications areas where DCT is used. In this paper we try to explore the utilization of discrete cosine transformation (DCT) domain with varying fractions of coefficients *i.e.* Fractional DCT (FRDCT) for feature extraction thereby intending to use spatial domain as well as frequency domain information simultaneously for face classification.

However, in case of automatic face recognition, the accurate recognition is heavily dependent upon the accuracy with which the features are extracted for face pattern representation as well as the performance of classifier used to distinguish between faces [1]. For the extraction and classification steps of image database *e.g.*: for the database that contains forty persons and ten instances of face image each of 512x512, the processing for single instance recognition requires processing of data of dimension $512 \times 512 \times 40 \times 10 = 104857600$. Due to the high dimension of data processing, the size of input data needs to be reduced first in order to reduce the number of computations and processing speed.

In this paper, we are using fractional discrete cosine transformation (FRDCT) definition [2] for feature representation of face image database whose dimensionality is further reduced by classical dimensionality reduction approach of Principal component analysis (PCA) [4]. The detail of feature selection and dimensionality reduction process followed is discussed in Section 3 where Section 2 deals with the related literature survey. Section 4 analyzes the experimental results of using FRDCT features of the input face images on the error rate in correct classification with BPNN classifier as well as simple Nearest Neighbour classifier. The experimental results discussed in section clearly shows that features extracted by using FRDCT have better classification accuracy with neural networks as classifier rather than the features extracted by the use of DCT. Section 5 concludes the paper with the discussion of impact on the classification error obtained by the virtue of the order parameter of the FRDCT. Section 5 also discusses the further scope which can further improve the cited results.

2. RELATED WORK

For reliable classification of pattern, the selection of feature space representing the pattern play very important role. The selection of appropriate feature space is desired as all the extracted features may not contribute to the classification positively. The presence of redundant features in pattern affects adversely the accuracy of classification [11]. For reliable recognition, it is desirable to extract features space

with accuracy from the training images even though some sort of distortion or deformation (up to certain extent) is present.

Discrete Fourier Transformation representation allows manipulating the information content of the image globally by means of spatial manipulations of the transformed image. The DFT representation also encodes image information in a holistic distributed manner [18]. However due to the fact that Discrete Cosine Transformation representation tends to have more of its energy concentrated in a small number of coefficients compared to DFT, it is actively used in spectral analysis [13]. Fractional version of cosine transformation has close relationship with conventional cosine transformation. Fractional cosine transformation provides a tool to integrate the information of time domain and frequency domain at the same time [9].

2.1 FRDCT (Fractional Discrete Cosine Transformation)

Fractional DCT is generalized form of conventional DCT with an additional free parameter and using this free parameter FRDCT makes its applications in all areas where DCT is used. Fractional version of Discrete Cosine Transform (FRDCT) has been used recently for feature extraction step of face recognition [19]. The feature selection process using FRDCT may involve either taking the transformation of the image as a whole or performing transformation block wise. The relevant coefficients are then separated. Mathematically, the 2D-FRDCT of an image with size dimensions [R, C] having R rows and C columns, with fractional order α and β is given by:

$$X_{(\alpha,\beta)}(m,n) = \sum_{p=0}^{R-1} \sum_{q=0}^{C-1} X(p,q) \text{Krn}l_{(\alpha,\beta)}(p,q,m,n) \dots \dots 1$$

The transformation kernel ($\text{Krn}l(\alpha, \beta)$) of Fractional DCT is defined by determining the fractional powers of the Eigen values of DCT kernel first row wise then column wise. The eigen values leads to a complex valued FRDCT matrix when $N \leq 4N_0$ i.e. number of input points is a multiple of four. Eigen Vectors and Eigen values help in finding the important attributes of the input image as Eigen values are useful for explaining the variability of pixel's intensity values. Only the components that exhibit the maximum variability play important role are kept and rest can be neglected. For N point DCT, the kernel is defined as

$$\text{Krn}l = \sqrt{\frac{2}{N}} * K_m * \cos\left(\frac{(2m+1)n\pi}{2N}\right) \dots \dots 2$$

where m, n = 1, 2, ..., N

$$K_m = \begin{cases} \frac{1}{\sqrt{2}}, & \text{if } m = 0 \text{ and } m = N \\ 1 & \text{for other values of m or n} \end{cases}$$

The fractional DCT is defined through the real powers of DCT Kernel matrix. The fundamental properties of Krnl are: (a) Krnl is real and (b) Krnl is orthogonal. The eigen decomposition of Kernel matrix can be expressed as:

$$(\text{Krn}l)^a = U * (D)^a * U^T \dots \dots \dots 3$$

where U is unitary matrix with columns of eigen vectors u_n , D is diagonal matrix with diagonal entries of eigen values λ_n raised by their a th power. A block diagonal matrix $(D)^a$ is computed by

$$D^a = \begin{bmatrix} G1_{N/2}(\theta(a)) & 0 \\ 0 & G2_{N/2}(\eta(a)) \end{bmatrix} \dots \dots \dots 4, \text{ where } G1 \text{ and } G2 \text{ are block diagonal matrix defined by: } G1_{N/2}(\theta(a)) = \begin{bmatrix} \cos(\theta(a)) & \sin(\theta(a)) \\ -\sin(\theta(a)) & \cos(\theta(a)) \end{bmatrix} \dots \dots \dots 5$$

The same way the G2 block diagonal matrix is defined with angle η . Among the choices for θ and η , we consider here $\theta(a) = 2\pi a$ and $\eta(a) = \pi a$, where a is the fractional order of the transform in time frequency plane. The value of a varies between 0 & 1 both inclusive i.e. $0 < a < 1$. For this range of a, FRDCT is partially like the Cosine transform and partially like the identity operation. When $a=1$, the FRDCT becomes the Discrete Cosine transform. When $a=0$, the FRDCT becomes the identity operation.

2.2 Dimensionality Reduction

Principal component analysis approaches the face recognition problem by means of information theory concepts [17]. The most relevant information in a face image is extracted and encoded as efficiently as possible. Recognition is performed on a face database that consists of models encoded similarly. In mathematical terms, the principal components of the distribution of faces or the eigenvectors of the covariance matrix of the set of face images, treating an image as a point (vector) in a very high dimensional face space is sought. These eigenvectors are called eigenfaces due to their resemblance of face images. Recognition is performed by obtaining feature vectors from the eigenvectors space [16]. PCA maximizes the variance of the projected samples. Principal Component Analysis [16,17] is defined by the transformation: $x_i = W^T * y_i \dots \dots \dots 6$, where $x_i \in X \subseteq \mathbb{R}^d$ and W is a d-dimensional transformation matrix whose columns are the eigenvectors related to the eigen values.

2.3 Feature Classification using Neural Network

In different fields of application including pattern recognition, identification and classification, neural networks have been trained to perform complex functions. Neural networks have the advantage that before testing with unknown instances, they can be trained to capture knowledge about the variation of face patterns through training images and thereby achieving good generalization [22]. For face recognition, which is inherently a nonlinear classification problem, different variants of ANN based classifiers are used for face recognition.

3. PROPOSED APPROACH

The approach proposed here use Fractional DCT with different rotation orders to transform image feature vectors into frequency feature vectors. The dimensionality of transformed feature set is further reduced by PCA approach. The reduced dimension set is given as input to the BPNN neural network for input face classification. The recognition rates, with varying dimensionality reduction factor and hidden layer neurons in neural network system are implemented. For evaluation purpose the proposed approach will be tested on the benchmark AT & T face database [15]. The database contains forty subjects and each subject has ten images. In the following experiments, five images of each individual from AT&T database are selected randomly for training set and the five face images for the test set. Below figure (Fig. 1) illustrates the process flow of the proposed approach.

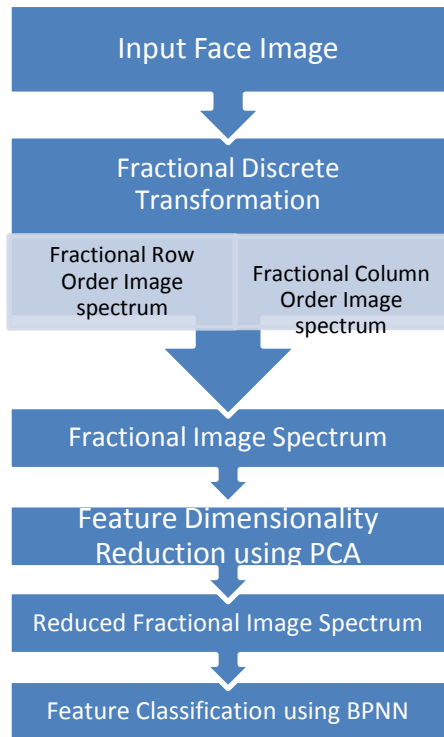


Fig. 1 (Proposed Approach)

3.1 Attribute Selection

The process of attribute selection extracts the most relevant information to be used to discriminating a person, i.e. the information that is common to a person and is different for the other persons [1]. At this stage of processing, from the input face of training database, the list of characteristics that best represents the face are selected and hence forms attribute list. Here in this paper we perform the attribute extraction, the approach using FRDCT frequency coefficients (Eq. 1-5). This approach is extension of fast and simple DCT. At the fractional order (α) which in our case it is 0.525, the frequency coefficients of 2D signal (image) is extracted. The frequency coefficients can be selected block by block or from the complete image. Here we have extracted the frequency coefficients from the complete image in order to reduce the computation complexity. Fig.3 shows the image retrieved back by using attributes represented by fractional order 0.5 – 0.7 followed by a normalization stage. As it can be seen, Mean square error increases as the alpha (rotation order) increases i.e. the image retrieved is becoming hazy at the higher order of rotation.

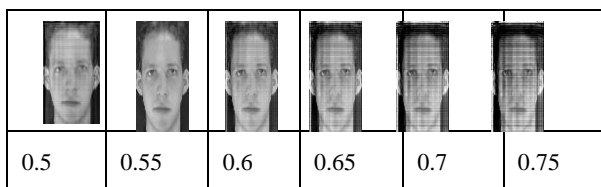


Fig. 2 (Image Retrieved at various Fractional orders)

3.2 Dimension Reduction using PCA

The high dimensionality of the related “image space” is a well-known problem for the design of a good verification algorithm [17]. The extracted attribute feature represents a d -dimensional feature vector where d is equal to the number of pixel of each face image. Processing d pixel values for N steps requires $N \times d$ computations. Therefore, PCA approach is used for reducing the dimensionality of the image space [16, 17]. PCA considers only a small number of “principal” components that exhibits the highest variance i.e. the most expressive features.

3.3 Training Stage

Neural networks have the advantage that before testing with unknown instances, they can be trained to capture knowledge about the variation of face patterns through training images and thereby achieving good generalization. The training stage performs various attribute selection steps and decides which one produces the best result [22]. In this paper, first five faces from AT & T database [15] were selected to train the network. The parameters of feature classifier are adjusted by using training sets. After doing this, for a given test image, face image is classified as either as one corresponding to the subject already existing in the database or to the unknown subject. To test the proposed method the remaining images of one subject are used. AT&T database images are in gray scale, with 256 different levels and a dimension of 92x112 pixels. It contains facial photos of 40 people with each person/subject photographed in 10 different poses, totalizing 400 photos (Fig. 1). The faces contain illumination variations, facial expressions variations (eyes closed/opened, smiling/not smiling), facial poses and facial details (with/without glasses, with/without beard), among other types of variations.

3.4 Testing Stage

The testing or classification stage compares the given input target face with a series of faces from testing database. The reduced set of FRDCT coefficients of test face is tested for similarity match with the coefficients of the train database. Here, the comparison is done using a similarity measure that indicates the resemblance between the test face and all the other faces by using BPNN as well as simple Nearest Neighbour Classifier that uses Euclidean distance as the matching criteria. The number of inputs to the neural classifier in this paper depends upon the number of principal components retained to represent the input face image.

4. EXPERIMENTAL RESULTS ANALYSIS

Experimental tests were conducted to validate the proposed method. This section presents the experimental tests that were carried out.

4.1 Error Rate vs. Order of Rotation

This experiment shows the variation of classification error rate when classified by the neural network classifier or by nearest neighbour clustering approach. In case of neural network classifier 145 hidden neurons were used at the hidden layer and forty principal components are retained from the 112x92 input FRDCT frequency coefficients. Following graph (Fig. 4) shows the variation of the correct classification error rate with the fractional order of rotation while obtaining fractional DCT attributes from the input image.

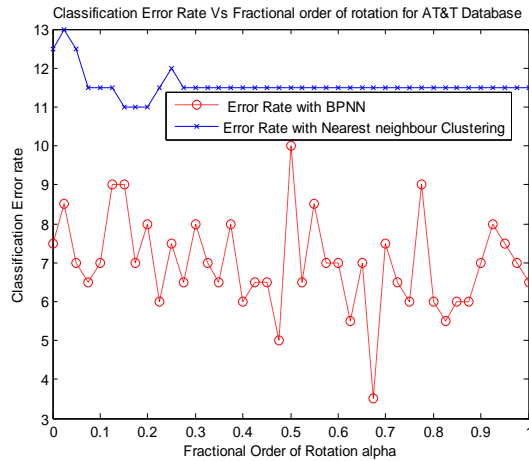


Fig. 3 (Classification error (%) for various orders of fraction (α))

From the experiment it was found that fraction order more than 0.6 and less than 0.7 gives lowest misclassification rate when classified with the Gradient descent training function of back propagation neural network. A steep fall in error rate is observed when classified with BPNN as compared to simple Euclidean distance classifier (Table 1).

Table 1 (Classification Error of FRDCT coefficients with Nearest Neighbor classifier* and BPNN **)

ErrFDCTPCA(*)	ErrNFDCTPCA(**)	Alpha
12.5	7.5	0
13	8.5	0.0250
12.5	7	0.0500
11.5	6.5	0.0750
11.5	7	0.1000
11.5	9	0.1250
11	9	0.1500
11	7	0.1750
11	8	0.2000
11.5	6	0.2250
12	7.5	0.2500
11.5	6.5	0.2750
11.5	8	0.3000
11.5	7	0.3250
11.5	6.5	0.3500
11.5	8	0.3750

11.5	6	0.4000
11.5	6.5	0.4250
11.5	6.5	0.4500
11.5	5	0.4750
11.5	10	0.5000
11.5	6.5	0.5250
11.5	8.5	0.5500
11.5	7	0.5750
11.5	7	0.6000
11.5	5.5	0.6250
11.5	7	0.6500
11.5	3.5	0.6750
11.5	7.5	0.7000
11.5	6.5	0.7250
11.5	6	0.7500
11.5	9	0.7750
11.5	6	0.8000
11.5	5.5	0.8250
11.5	6	0.8500
11.5	6	0.8750
11.5	7	0.9000
11.5	8	0.9250
11.5	7.5	0.9500
11.5	7	0.9750
11.5	6.5	1.0000

4.2 Error Rate vs. Hidden Number of Neurons

This experiment is done to find the optimum number of hidden number of neurons required for the neural training network so that testing error can be minimized. From the experiment it is found that with five training images and with rotation order 0.525, 145 neurons are sufficient to produce satisfactory recognition rates. However this large number of hidden neurons increases computation load .Sixty five hidden neurons also generates error rate near to that of obtained by 145 hidden neurons.

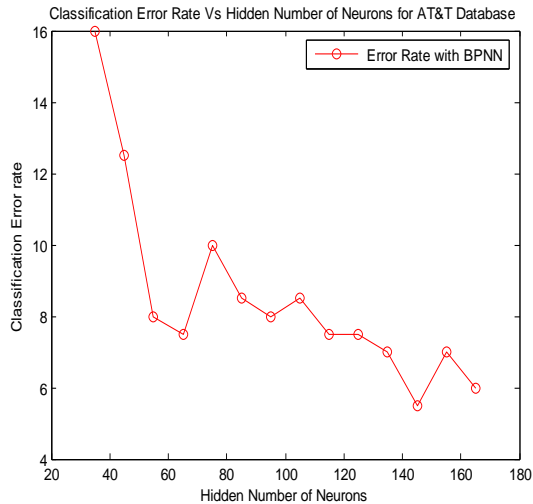


Fig. 4 (Classification error rate for different number of hidden neurons)

4.3 Error Rate vs. Number of Retained Principal Components

Number of principal components retained during application of dimension reduction PCA approach has significant effect on the convergence time of the neural network classifier. With less number of components, the number of epochs required during network testing reduces significantly. The following graph shows the network classification error rate during testing phase.

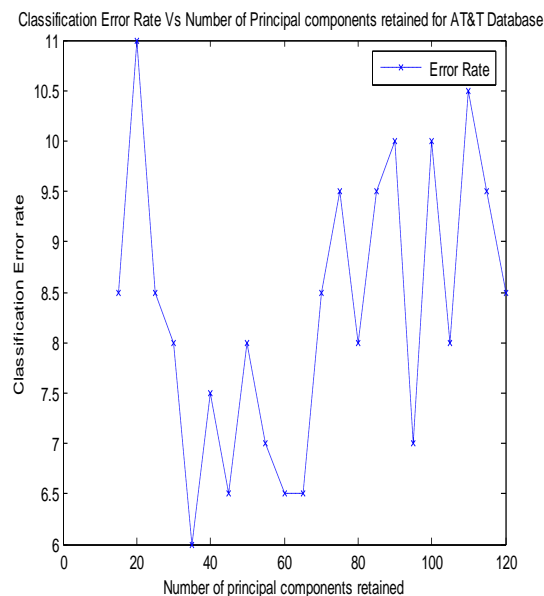


Fig. 5 (Classification error for different Principal Components retained)

It can be observed that there is sharp variation in the correct classification as the number of retained components increases. The sharp variations can be accounted for the noise and feature variations among the subject's instances present in the (a) Transformed image (b) Original image (c) Reduced image. As the number of retained components increases the efficiency is dropping due to the dimensionality curse problem (Table 2).

Table 2 (Error Rates obtained with different Classifiers for different Classifiers)

ErrFDCTPCA (Nearest Neighbour Classification)	ErrNFDCTPCA (BPNN Classification)	Principal components retained
13.5000	8.5000	15
11.5000	11.0000	20
14.0000	8.5000	25
11.5000	8.0000	30
12.0000	6.0000	35
11.5000	5.5000	40
11.5000	5.5000	45
11.7950	8.0000	50
12.0000	7.0000	55
13.0000	6.5000	60
13.5000	6.5000	65
14.0000	8.5000	70
14.0000	9.5000	75
15.0000	8.0000	80
14.0000	9.5000	85
14.0000	10.0000	90
14.5000	7.0000	95
15.0000	10.0000	100
15.5000	8.0000	105
17.0000	10.5000	110
15.5000	9.5000	115
16.5000	8.5000	120

4.4 Error Rate vs. Number of Training Images

In this experiment, error rate is studied versus the number of images used for training. From the given graph (Fig. 6) it can be observed that the error rate in correct classification drops rapidly with the increase in number of images. For large number of training images the classification accuracy of both the classifiers (nearest neighbour classifier and neural network classifier) approaches the same.

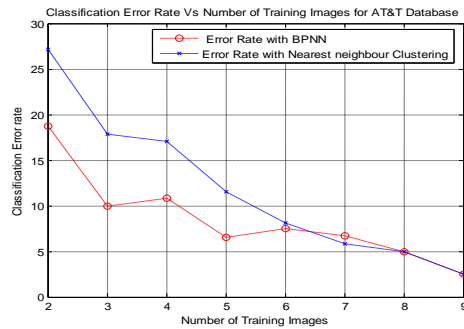


Fig. 6 (Classification error rate for different number of training images)

5. CONCLUSION AND FUTURE WORK

In this paper, facial features are extracted by the global application of FRDCT. Data dimensionality of extracted feature set is reduced by PCA. For AT&T database, it is found that for the whole database, 5 training images per person, 40 eigen face for total database and 40 neurons in hidden layer are enough for satisfactory recognition rates (about 92 %). For the classification purpose BPNN is used even though the learning speed of this type of training algorithms becomes sometime very slow. In this paper the convergence speed is increased by retaining limited number of principal components representing the complete dimension of image set of the database. Significant improvement is correct classification of images when BPNN is used as classifier over the nearest neighbour classifier is cited in this paper. For future work we would like to explore effect of the classification rate with the non-linear neural network classifier like Radial Basis function Neural Network or General Regression Neural Network.

6. REFERENCES

- [1] Li S.Z. and Jain A.K. 2005. Handbook of Face Recognition, Springer, New York.
- [2] Cariolaro G., Erseghe T. and Kraniuskas P. 2002. The Fractional Discrete Cosine Transform. IEEE Trans. on Signal Proc., 50(4). (902-1002).
- [3] S-C. Pei and M-H. Yeh. 2001. The Discrete Fractional Cosine and Sine Transforms. IEEE Trans. on Signal Proc. 49(6). (1198-1207).
- [4] Candan C., Kutay M.A. and Ozaktas H.M. 2000. The Discrete Fractional Fourier Transform. IEEE Trans. on Signal Proc., 48(5). (1329-1337).
- [5] Saxena R. and Singh K. 2005. Fractional Fourier transform- A novel tool for signal processing. J. Indian Inst. Sci., 85. (11-26).
- [6] Pei S.C. 2006. The multiple-parameter discrete fractional Fourier transform. IEEE Signal Processing Letters. 13(6). (329-332).
- [7] Hafeed Z. M., Levine M. D. 2001. Face recognition using the discrete cosine transform. Intl. Journal of Computer Vision. 43(3). (167-188).
- [8] Ozaktas H.M, Zalevsky Z. and Kutay M.A. 2000. The fractional fourier transform with applications in optics and signal processing. Wiley, New York .
- [9] Almeida L.B. 1994. The fractional Fourier transform and time-frequency representations. IEEE Trans. on Signal Proc. 42. (3084-3091).
- [10] Ozaktas H. M. , Arikan O., Kutay M.A. and Bozdagi G. 1996. Digital computation of the fractional Fourier transform. IEEE Trans. on Signal Proc. 44. (2141-2150).
- [11] Namias V. 1980. The fractional order Fourier transform and its applications to quantum mechanics. J. Inst. Math Applications. 25. (241-265).
- [12] Tsai H-H. 2007. Decision-Based Hybrid Image watermarking in wavelet Domain Using HVS and Neural Networks. In D Liu et al (Eds.) ISNN 2007. LNCS-4493. (904-913). Springer, Heidelberg.
- [13] Khayam S.A. 2003. The Discrete Cosine Transform (DCT): Theory and Application. DCT-Tutorial. WAVES-TR-ECE802.602.
- [14] Lam E.Y. and Goodman J.W. 2010. A Mathematical Analysis of the DCT Coefficient Distributions for Images. IEEE Trans. on Image Proc. 9(10). (1661-1666).
- [15] www.face-rec.org
- [16] Kirby M., and Sirovich L., "Application of the Karhunen-Loeve procedure for the characterization of human faces", IEEE PAMI, Vol. 12, pp. 103-108, (1990).
- [17] Sirovich L. and Kirby M. 1987. Low-dimensional procedure for the characterization of human faces. J. Optical Society of America, 4(3). (519-524)
- [18] Jain. A.K. 1989. Fundamentals of Digital Image Processing. Englewood Cliffs, NJ: Prentice Hall.
- [19] Shekhar B.H., Thippeswamy G. and Kumari.S. 2011. Face Recognition Based on Fractional Discrete Cosine Transform. IEEE Int. Conf. on Recent Trends in Information Technology. (987-991).
- [20] Singh N. 2013. M.Tech Dissertation. Image Compression and Encryption Using Discrete Fractional Cosine Transform. THAPAR Univ., India.
- [21] Jindal N. and Singh K. 2013. Video Compression-Encryption using Three Dimensional Discrete Fractional Transforms. Research J. of Applied Sciences, Engineering and Technology. 5(14). (3678-3683).
- [22] Haykin S. 1999. Neural Networks- A Comprehensive Foundation, 2nd Ed. Prentice Hall.