iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 5 **Issue:** IX **Month of publication:** September 2017

DOI:

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor:6.887
Volume 5 Issue IX, September 2017- Available at www.ijraset.com

A New Paradigm of Face Recognition Using Multi-class Classifier

Deepali H. Shah[1], Dr. J. S. Shah[2]
[1]Research Scholar, School of Engineering, R. K. University, Rajkot, Gujarat, India
[1]Associate Professor, Instrumentation & Control Department, l. D. College of Engineering, Ahmedabad, Gujarat, India
[2]Ex-Principal, Government Engineering College, Patan, Gujarat, India

Abstract: ***Eminent applications for security and law enforcement have enhanced significant importance of automatic face recognition. Face recognition is one of the most convenient ways for identification and recognition of a person in considerable large database. The face recognition complications are due to change in illumination of light sources, variations in pose as well as facial expressions. This paper will present a novel approach of hybrid face recognition technique. The face recognition system is developed with PCA and SVM as multi class classifier. The experiments of the proposed method have been conducted on ORL data sets. The experiment result is presented and discussed, which shows the effectiveness of the strategy described.***
Keywords: ***PCA, SVM, eigenfaces, OSH, ORL***

I. INTRODUCTION

Due to significant rise in terrorist activities, importance of more sophisticated security system for identification of individuals has increased rapidly. Many researchers have devoted their works in field of biometric and non-biometric methods for identification of individuals. Systems like ID numbers, different passwords, ID cards are used to represent a person in non-biometric methods. The accuracy of non-biometric systems is challengeable due to hackers' increasing ability to hack password and many more ways to forge the same. On the other end, unique observable characteristics of individual are used by biometric recognition methods to identify a person. Many researchers have developed biometric recognition systems based on unique features of individuals like voice, iris, palm geometry, finger print, signature as well as facial features for identification of person which furnish highly accurate recognition systems[1]-[2]. Although, biometric recognition systems like finger print scanning and iris identification are practically liable to error due to intrusive identification. Moreover, essential co-operation required from users for identification and highly expensive equipments lead them towards less acceptability [1]-[2]. On the contrary, the face recognition system receives considerable attention as cameras at surveillance place captures images of all the objects in real time, continuous in nature and most importantly without anybody's notice [2].

Many researchers have contributed their work in face detection and recognition in recent years [1]-[2]. Most of the researchers' work is based on certain preconditions which lead to many limitations. As a result of these limitations and performance issues, practical implementation in real environment is quite less [2]. In this research, novel approach of hybrid face recognition technique is introduced. The face recognition system is developed with PCA and SVM and will be checked with ORL data sets. Face recognition comprises of three major steps which are detection, feature extraction, and classification. Optical image captured by sensor is converted into digital image in detection phase of face recognition. Features are extracted from faces in second step of feature extraction. Data are arranged in groups or class as per the similar characteristic in third step of classification [3]. PCA technique is used for feature extraction which emphasis on the problem of scalability. The popular method to solve structural risk minimization and best ever compromise between learning accuracy and learning ability, under the finite sample circumstances is adopted in SVM. The classification step is carried out by SVM in which generalize performance is achieved [3]. The settlement of best kernel and choose of optimal feature enhances the performance of face recognition system which leads to combination of PCA and SVM for implementation [3].

II. PRINCIPAL COMPONENT ANALYSIS

Principal component is extracted mathematically from multi-dimensional data and dimensional reduction is carried out in PCA. The highest variability resides in the linear combination of the original dimensions. It is the first principal component in PCA. The maximum variability resides in the linear combination of n-th principal component which is being orthogonal to n-1 first principal components [4].

The face images are decomposed in to small set of characteristic feature images which is called 'Eigenfaces'. This Eigenfaces are initial training sets comprising principal components. The PCA scheme use these informations based on an information theory approach. The original image can be reconstructed from training set by combination of Eigenfaces. This is a very important feature of PCA. The eigenfaces are created and recognition is carried out by them. This technique is also outside of facial recognition [5]. The feature is transformed in mathematical sense from physical face by mathematical transform for recognition in terms of eigenface [6].

The face recognition using eigenfaces comprises two phases. The training phase is the first one of that. In the first phase, individual faces comprised in large group are acted as a training set. These individual face training images must be good representation of all possible faces that one might encounter [6].

Column vector is represented by each face image with each entry corresponding to an image pixel during training phase. The normalization with respect to average face is carried out for these image vectors. Speedup technique is used to find the eigenvectors of the covariance matrix of normalized faces in the algorithm. Number of multiplications to be performed is reduced due to speedup technique. Successively, the eigenvector matrix is multiplied by each of face vectors to obtain their corresponding face space projections. Maximum distance between any two face projections is used to compute the recognition threshold at the end [7]-[8].

Recognition phase is the second phase of the algorithm. In the second phase, extraction of new image is carried out. The image is subtracted from the average face ψ for recognition of the image. The dot product of the input vectors with the eigenfaces is carried out. Thus, face space is elaborated with projection of the input image. The face space is also elaborated with projections of the training image in similar way. The computation for euclidean distances of point of the input image with the points of training set is carried out subsequently. The best match is the training set image with minimum distance from the input image [6]. The comparison of smallest distance with threshold θ decides whether face is known or not. If the obtained distance is greater than θ, the face is considered as new one else the face match is considered [9].

III. SUPPORT VECTOR MACHINE

The two data sets with maximum distance between them are separated in a classification method which is called as SVM. SVM was proposed by Vapnik in 1998 [10]. The hyperplane is found in SVM in which the separation of the largest possible fraction of points of the same class on the same side is obtained and maximization of the distance from either class to the hyperplane is carried out [11].

Separation of two data sets is carried out by searching for an optimal separating hyperplane (OSH) between them in SVM. ''Support vectors" are the bounds between data sets and OSH [10].

The data, that are not linearly separable, are transformed into new space using kernel and OSH is found. Maximization of the margin and minimization of the misclassifications is evaluated if data is separable and OSH is searched. The vector $\boldsymbol{w}$ is nomenclature as the normal vector, x as an input data point vector and b is a constant [12].

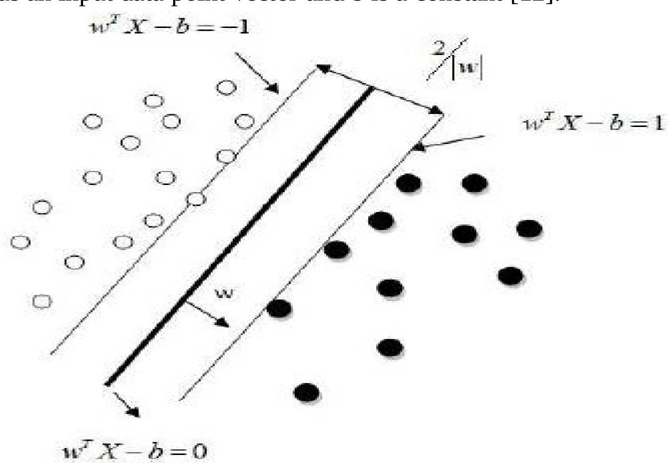


Fig.1 Support vectors in SVM [13]

Each point of total data set is referred as $x_i \in IR^n$, i = 1, 2. . . N and belongs to a class $y_i \in \{-1, 1\}$. Identification of two classes and the OSH separating them for linear classification are like,

$$w \,.\, x_i + b \geq 1, \quad y_i = 1 \tag{1}$$
$$w \,.\, x_i + b \leq -1, \quad y_i = -1 \tag{2}$$

The generalized form may be,

$$y_i \,.\, [w \,.\, x_i + b] \geq 1, \quad i = 1, \ldots, l \tag{3}$$

The pre-defined distance between support vectors is as:

$$d = \frac{2}{\|w\|} \tag{4}$$

The better separation between two classes can be achieved if bigger d is obtained. Maximization of d can be elaborated with minimization norm of w for this reason. Lagrange function is used to solve this problem as stated below

$$L(w, b, \alpha) = \frac{\|w\|^2}{2} - \sum_{i=1}^{l} \alpha_i \,.\, \{y_i \,.\, [w \,.\, x_i + b] - 1\} \tag{5}$$

where, α_i represents Lagrange multipliers. Solve (5) by minimizing according to w and b as well as maximizing according to $\alpha_i \geq 0$ values, most suitable OSH parameter w can be obtained in (6) according to condition

$$\sum_{i=1}^{l} \alpha_i \,.\, y_i = 0, \alpha_i \geq 0, i = 1, \ldots, l$$
$$w = \sum_{i=1}^{l} \alpha_i \,.\, y_i . x_i, , \alpha_i \geq 0, i = 1, \ldots, l \tag{6}$$

Distance of any data point x to OSH is defined as

$$d(w, b, x) = \frac{|w.x + b|}{\|w\|} \tag{7}$$

More generalized form of (7) can be obtained by replacing w with its value in (6),

$$d(x) = \frac{\left(\sum_{i=1}^{l} \alpha_i \,.\, y_i . x_i\right) . x + b}{\left\|\sum_{i=1}^{l} \alpha_i \,.\, y_i . x_i\right\|} \tag{8}$$

Sign of distance calculated in (8) confirms to which class point x belongs and | d | shows distance of x to OSH. Better classification result can be obtained by increase in | d | [10].

Linear separation of data sets is not being achievable successfully all the time. Simple conversion of feature space is to be carried out in such cases. Point x in first data space is expanded to a feature space with higher dimension and linear separation is retried. This expansion process is realized with operator $\emptyset(.)$ OSH function turns into the form:

$$f(x) = w . \emptyset(x) + b$$

By replacing w with its value in (6), the more generalized form is obtained as:

$$f(x) = \sum_{i=1}^{l} \alpha_i \,.\, y_i \left(\emptyset(x_i) . \emptyset(x)\right) + b$$

The ‘‘Kernel Functions” in $K(x_i, x) = \left(\emptyset(x_i) . \emptyset(x)\right)$ form are used in a high dimensional space realization of $\left(\emptyset(x_i) . \emptyset(x)\right)$, where multiplication is intractable [10].

The ``one-versus-all" or OVA classification is the most frequent method for multiclass classifier. It is used to construct one-vrsus-rest classifiers where each category is split out and all of the other categories are merged as well as to choose the class which classifies the test data with greatest margin. Thus, m class problem is divided into m binary problems. The whole training data is used for learning step of the classifiers. The patterns from the particular class are considered positives while all other examples are considered negatives [14].

The pattern is presented to each one of the binary classifiers in validation phase. The classifier with positive output indicates the output class. Some tie-breaking techniques are quite compulsory as the positive outcome is not unique in numerous cases. Predicting the class from the classifier with the maximum confidence is the most familiar approach which uses the confidence of the classifiers and decides the last outcome. The outcome of OVA classifiers (where r_i in [0, 1] is the confidence for class i) is a score vector instead of score matrix:

$$R = (r_1, r_2, \ldots\ldots r_i, \ldots\ldots r_m)$$

The m class problem is divided into $m(m-1)/2$ binary problems for construction of a set of one versus-one classifiers by another strategy. Each problem emerges up by a binary classifier and it is responsible of distinguishing between a different pair of classes.

The training data, which are a sub set of instances from the original data-set and contain any of the two corresponding class labels, is used for the learning phase of the classifier. The instances with dissimilar class labels are ignored at that time [14].

The pattern is presented to each one of the binary classifiers in validation phase. The output of a classifier given by r_{ij} in [0, 1] is the confidence of the binary classifier discriminating classes i and j in favour of the former class. The output class of a classifier is a class with the largest confidence. The said outputs are represented as a score matrix R:

$$R = \begin{pmatrix} - & r_{12} & \cdots & r_{1m} \\ r_{21} & - & \cdots & r_{2m} \\ \vdots & \vdots & \cdots & \vdots \\ r_{m1} & r_{m2} & \cdots & - \end{pmatrix}$$

IV. RESULTS AND ANALYSIS

To test the performance, some experiments are performed on a set of face images taken at the Olivetti Research Laboratory (ORL) in Cambridge University, U.K, which can be used freely for academic research. ORL face database contains images of 40 distinct persons, each person having ten different face images. There are 400 face images in total, with 256 gray degrees and the resolution of 92 x 112. For some subjects, the images were taken at different times, which contain quite a high degree of variability in lighting, facial expression (open/closed eyes, smiling/non-smiling etc), pose (upright, frontal position etc), and facial details (glasses/no glasses). All the images were taken against a dark homogeneous background with the subjects in an upright, frontal position, with tolerance for some tilting and rotation of up to 20°. The variation in scale is up to about 10%.

Here, we are using PCA to train the faces. Then SVM is used on training data along with the labels associated with each training face. During recognition phase we use support vector machine as a classifier. In the recognition phase, first PCA is used and then SVM is used to classify a test sample using trained SVM. It retrieves a set of support vectors and calculates the total match to recognize a test face.

The experiments were carried out on 100, 200, 300 and 400 test images of 10, 20, 30 and 40 persons respectively using PCA with SVM.

Table I shows performance accuracy with varied no. of testing faces. The number of persons is varied in training sets and testing sets. In training set, 1 image per person is used. In testing set, 10 images of each person are used.

TABLE I
PERFORMANCE BASED ON VARIED NUMBER OF FACES FOR MODIFIED PCA WITH SVM

Training Faces 1 face/person	Testing Faces Total Faces	Performance %
40	400	76.75%
30	300	84.7%
20	200	83.5%
10	100	90%

Table I shows that as the no. of persons increase in training faces, the accuracy decreases. This is shown in the graph of Fig. 2.

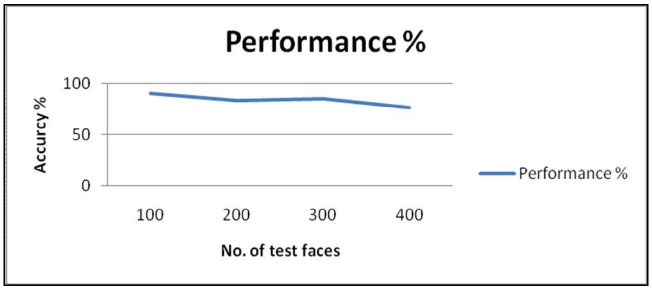


Fig. 2 Performance Chart

Experiments were carried on PCA for different number of persons. Then same experiment was carried out by modifying the PCA. In modified PCA, training set is internally divided into two parts and average face of each part is calculated individually and then average face of two average faces is calculated. This gives fine refinement in average face which helps in increase in accuracy performance. Here, for SVM, constant hyperplane is used.

Then experiments were carried out on PCA with SVM. Here, SVM is used as a binary multi-class classifier.

Experiments were carried out on different sets of training faces and testing faces. In training set, 1 image per person is used while in testing, 10 images per are used. Experiments were carried out for 10, 20, 30 and 40 persons.

TABLE III

PERFORMANCE COMPARISON OF PCA, MODIFIED PCA WITH CONSTANT HYPERPLANE AND MODIFIED PCA WITH SVM

Training Faces 1 face/person	Testing Faces Total Faces	PCA Performance %	Modified PCA with constant hyperplane Performance %	Modified PCA with SVM Performance %
40	400	30%	74%	76.75%
30	300	40%	80.33%	84.7%
20	200	30%	78.5%	83.5%
10	100	50%	90%	90%

Table II shows the comparison of performance of PCA and modified PCA with two classes.

Fig. 3 shows that with PCA, accuracy obtained is 30-50%, while with modified PCA with constant hyperplane, accuracy obtained is 74-90% and with modified PCA with SVM, accuracy obtained is 76.75-90%

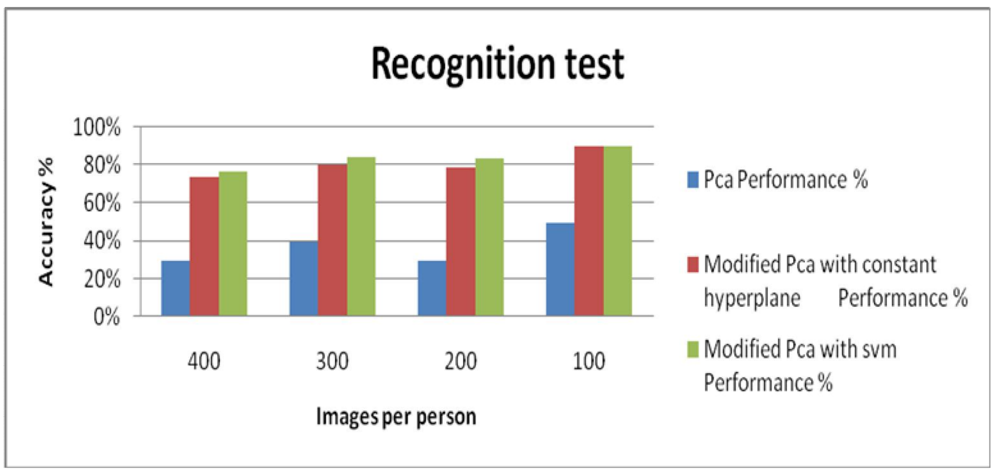


Fig. 3 Comparision Chart for PCA, modified PCA and PCA with SVM

V. CONCLUSIONS

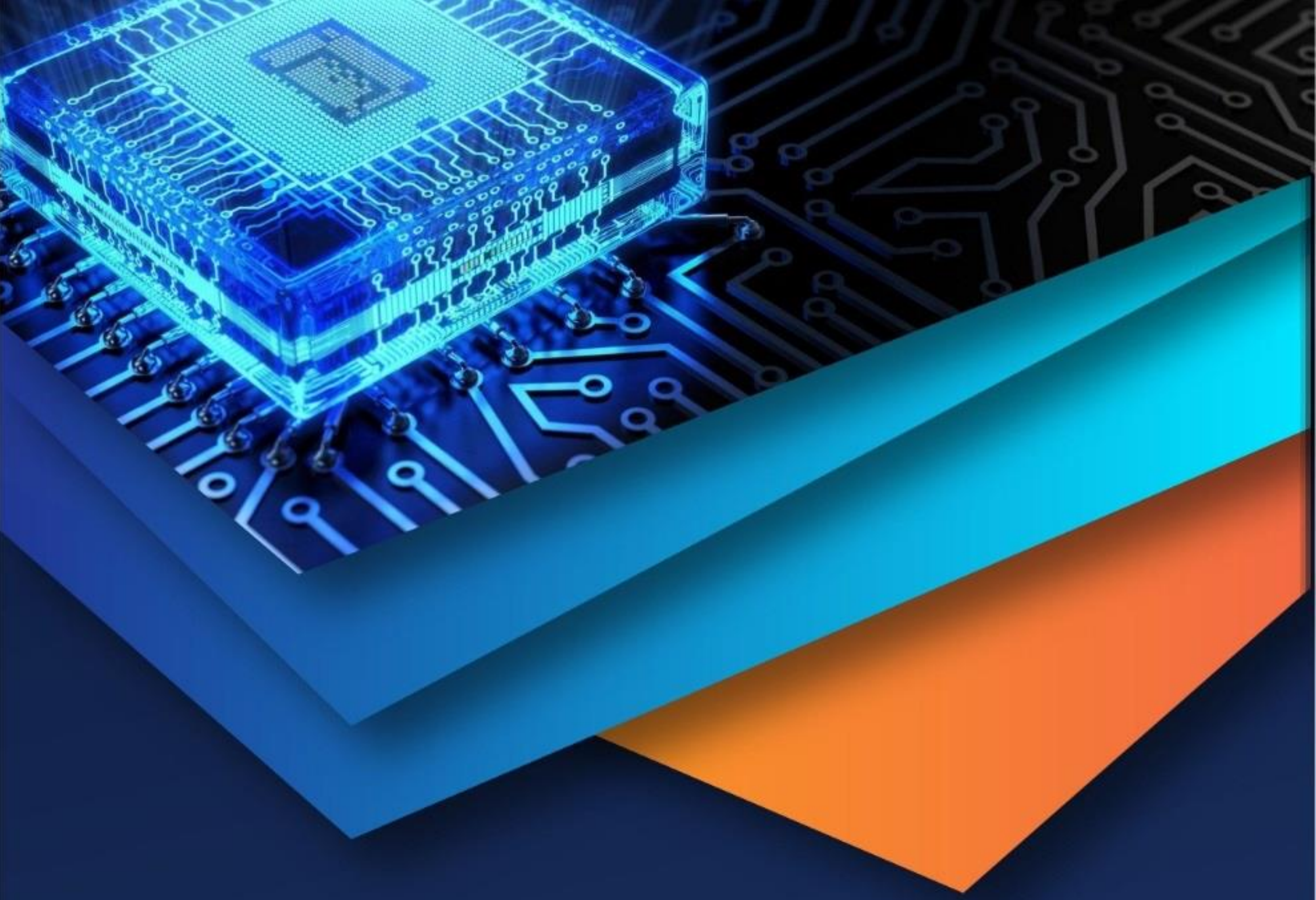
Face recognition method using eigenfaces is proposed. We used ORL database of face images which contains 400 images of 40 different persons, each individual have 10 images. From the results, it can be concluded that, PCA gives 30-50%. By using modified PCA with constant hyperplane, average image gets more refined and accuracy increased upto 74-90%. Accuracy of 76.75-90% is obtained with modified and combined PCA and SVM. This research is aimed to enhance the face recognition application performance in respect to present scenario. The proposed improved face recognition application may be quite useful in security systems, Process plant interface with working personal, Visa offices, virtual entertainment, banks or even in mobile and web security.

REFERENCES

[1] Andrea F. Abate, Michele Nappi, Daniel Riccio, Gabriele Sabatino (2007), "2D and 3D face recognition: A survey", 2007 Elsevier B.V. Available: http://codc4fv.googlecode.com/files/faceSurvey2007.pdf

[2] PAO-CHUNG, CHAO. FACE "RECOGNITION SYSTEM BASED ON FRONT-END FACIAL FEATURE EXTRACTION USING FPGA", M.Sc. Thesis, De La Salle University, Manila, 2009

[3] Rakesh Kumar Yadav, Dr. A. K. Sachan, Dr. D. Rai, "State of Art in Kernel Classifier SVM with PCA for face recognition: A Litrature review", Journal of Computer Engineering Vol 3, No 1 (2012), ISSN: 2010-1619, pp. 6-16
[4] PF de Carrera, "Face Recognition Algorithms", 2010, Avaiklable: www.ehu.es/ccwintco/uploads/e/eb/PFC-IonMarques.pdf
[5] Sanjeev Kumar and Harpreet Kaur, "FACE RECOGNITION TECHNIQUES: CLASSIFICATION AND COMPARISONS", International Journal of Information Technology and Knowledge Management July-December 2012, Volume 5, No. 2, pp. 361-363
[6] Tang Ho Man, Sunny. "Face Recognition Algorithms Review Term Paper" - December 2001. The Chinese University of Hong Kong
[7] P. Kannan,Dr. R. Shantha Selva Kumari, "A Literature Survey on Face Recognition System: Algorithms and its Implementation", Available: http://share.pdfonline.com/af74b8894437414695fb5abce41a0fdb/A_Literature_Survey_on_Face_Recognition.htm
[8] M.A. Turk, A.P. Pentland, "Face Recognition Using Eigenfaces", IEEE Conference on Computer Vision and Pattern Recognition, pp.586--591, 1991.
[9] Madi, Lahoud et. al., "Face Recognition on FPGA"
[10] Ergun Gumus et. al., "Evaluation of face recognition techniques using PCA, wavelets and SVM", An international journal: Expert Systems with Applications 37 (2010) 6404–6408, Elsevier publication
[11] Saeed Meshgini, Ali Aghagolzadeh, and Hadi Seyedarabi, "Face Recognition Using Gabor Filter Bank, Kernel, Principle Component Analysis and Support Vector Machine", International Journal of Computer Theory and Engineering, Vol. 4, No. 5, October 2012, pages:767-771
[12] Manisha Satone, Gajanan Kharate, "Face Recognition Technique Using PCA, Wavelet and SVM", International Journal on Computer Science and Engineering (IJCSE), Vol. 6 No. 01 Jan 2014, ISSN : 0975-3397, pages-58-62
[13] Chandrakamal sinha, Thesis on "Gender Classification from Facial Images using PCA and SVM", June 2013
[14] Yashima Ahuja and Sumit Kumar Yadav, "Multiclass Classification and Support Vector Machine", Global Journal of Computer Science and Technology Interdisciplinary, Volume 12, Issue 11, Version 1.0, 2012.



10.22214/IJRASET

INDEX COPERNICUS
45.98

IMPACT FACTOR:
7.129

TOGETHER WE REACH THE GOAL
IMPACT FACTOR:
7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY