Performance Evaluation on KVKR- Face Database using Multi Algorithmic Multi Sensor Approach

Siddharth B. Dabhade
Dept. of Computer Science & IT
Dr. Babasaheb Ambedkar Marathwada University
Aurangabad

Nagsen S. Bansod
Dept. of Computer Science & IT
Dr. Babasaheb Ambedkar Marathwada University
Aurangabad

Yogesh S. Rode
Maharashtra University of Health Sciences
Nashik

M. M. Kazi
Dr. G. Y. P. College, MGM
Aurangabad

K. V. Kale
Dept. of Computer Science & IT
Dr. Babasaheb Ambedkar Marathwada University
Aurangabad

ABSTRACT
Biometric is emerging area in the computer science for the secure various systems. Day to day life peoples are preferred to use robust and highly acceptable security system which can surpass the human errors. Many scientists are engaged to develop strong biometric system but there are a lot of challenges in the real time application. It is observed and found that researchers are only working on too old laboratory databases such as ORL. But now a day's various cost effective data acquisition sensor are coming in the market with high resolution of data. When we are using different type of data capturing devices gives difference in performance of recognition rate. In this work we have proved that recognition rate is affected by the various sensor as well as database environment. For robust face recognition system suitable algorithms are suggested to different type of sensors.

General Terms
Pattern Recognition, Security, Biometrics, Image Processing, Signal Processing

Keywords
Face Recognition, KPCA, KFA

1. INTRODUCTION
The Biometric is a process of human authentication system through their physical, behavioral or chemical properties. Face, Fingerprint, Iris, Ear, etc. are comes the under the physical biometrics where as voice, gesture, signature, etc. are behavioral and DNA, Blood Group, etc. are chemical properties. Face recognition is a biometric trait used to identify the person on his physical look. By using face various aspects such as gender, age, ethnicity and various emotions of human beings. There are many applications of face such as in the border crossing, identify the missing children's, law enforcement, government & private organization, etc. There are various approaches are used to detect the face and identify the person on the basis of appearance, model, texture, etc. Face recognition is very difficult task while facial expression, mustaches, beard, wearing specs, over makeup, mask, jewelry, in case of twins and illuminations [1] [2]. On international level tremendous researchers are developing face recognition applications which could provide robust and secure system but still research is in the laboratory. Till date approximately 76 various database sources are available [3] apart from these we have selected widely used database i.e. ORL database and newly evolved KVKR- Face Database. This paper is divided mainly in 5 sections. First section gives the introduction to the face biometrics system. Fundamental concepts of face recognition are explained in the section two. Section three gives detail explanation of proposed system, feature extraction, classification, etc. Section 4 gives experimental results with graphical representation of major contribution of comparison. Last section gives overall conclusion of the paper.

2. FUNDAMENTAL CONCEPTS
Face recognition is a very interesting area in biometrics. Face recognition is an active research area for pattern recognition and computer vision. It is a difficult and complex problem and due to its potential use in a wide variety of business and law enforcement uses including access control, security monitoring, and video surveillance. An another biometric identification systems based on physical characteristics, face recognition is a passive, non- intrusive system for prove own identity in a user-friendly way without having to interruption [4]. It has many application areas, i.e. human computer interaction, security [5]. One of the factors that affect the performance of the recognition system is the training sample size [6][7][8][9]. Sufficient number of training samples are always needed to train the classification system well [10]. If only one image per person is available, the recognition process gets more difficult. This problem is called one sample problem [11] [12]. Traditional methods will suffer or fail when a single image per person is available [13][14][15]. Several algorithms have been proposed to overcome this difficulty [16] [17][18].

Face Recognition System divided in four Section i.e. Data Acquisition, Feature Extraction, classification and Decision. There are tremendous sensor are available 2D, 3D & Hyper Spectral to acquire the face images standalone as well as remotely. Feature extraction techniques mainly classified into Photometric and Geometric (Feature Based). The Photometric is view based technique is an again divided into sub cluster such as statistical, Neural Network, Hybrid approach, etc. Statistical, Linear, Non Linear, transformed based, PCA, RBF, Range Data, Infrared, Profile, Active Shape Model, Wavelets, Elastic Graph Matching, Local Feature Analysis, etc. Euclidian, MahCos, HMM, GMM, SVM, etc. are widely used for classification. The threshold value or rank level is used to take the decision of the system. Decision system is depending on the identification or verification. In identification system user's identity is matched with N no. of samples available in the database where as in the verification system one to one matching is performed. System performance measures play an important role for result analysis these are False Acceptance Rate (FAR), False Rejection Rate (FRR), Equal Error Rate (ERR), Cumulative Match Curve (CMC), Receiver Operating Curve (ROC), etc.

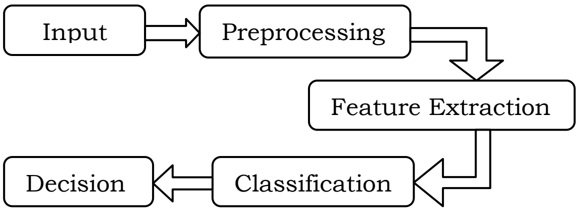

Fig.1 Generic Structure of Proposed Methodology

3. PROPOSED METHODOLOGY

Proposed methodology is consisting of input, preprocessing, feature extraction, classification & decision steps. Generic structure of proposes methodology is shown in Fig.1.

3.1 Input

In this experiment input is always face image which is from well known ORL database in gray scale and KVKR- Face Database in color images. Details of specification of both database is explained in the section number 4.

3.2 Preprocessing

These databases are contained an image format not scaled, after acquisition convert the color image into the gray level.

3.3 Feature Extraction

For feature extraction we have used various linear and non linear based statistical approaches are as follows:

3.3.1 KFA

By considering the set of image samples X_k,

$$x_k = [x_{k1}, \ldots, x_{kn}]^T \in R^n \tag{1}$$

KFA is performed using the similar procedure of KPCA except that Fisher Linear Discriminant (FLD) is considered instead of PCA after the transformation of the subspace to higher dimension. If xk has the same value of equation (1), the same projection is performed on the vector x. to get the function Φ: $R\ n \rightarrow R\ f$, f > n. Let the projected samples $\Phi(x)$ be centred in $R\ f$ and let the equations that use dot products be formulated for Fisher linear Discrimate Analysis (FLD) only.

Assume the within-class and between-class scatter matrices be $SW\ \Phi$ and $SB\ \Phi$, to apply FLD in kernel space, the solution to eigenvalues λ and eigenvectors

$$w\Phi \text{ of } SW\ \Phi w\Phi = SB\ \Phi w\Phi \tag{2}$$

are derived by finding the eigenvectors corresponding to largest generalized eigenvalue. The kernel function is introduce defined by

$$(\mathrm{krs})\mathrm{tu} = \mathrm{k}(\mathrm{xtr}, \mathrm{xus}) = \Phi(\mathrm{xtr}).\Phi(\mathrm{xus}) \tag{3}$$

where there exists a c-class problem and a r-th sample of class t and the s-th sample of class u be xtr and xus respectively (where class t has lt samples and class u has lu samples).

Then finally project $\Phi(x)$ to a lower dimensional space spanned by the eigenvectors $w\Phi$ in a way similar to Kernel PCA [19].

3.3.2 PCA

Linear and Non-Linear statistical techniques are used for reduced the dimensions from the given data set. One of the linear statistical powerful methods for extraction of structure from potentially high-dimensional data sets i.e. eigenvectors that are associated with the largest Eigen values from the input distribution, is a Principal Component Analysis (PCA). Principle Component is a set of observation of possibly correlated variables into a set of values of orthogonal transformation of linearly uncorrelated variables. Possibility of variance is less than or equal to the variable sets in the given data organized by the descending order i.e. first value is highest possible variance under the various constraint, the procedure of this uncorrelated arrangement of the variances is called as orthogonal transformation. If data is normally distributed it performs well. PCA is responsive to the relative scaling of the original variables. On the basis of various applications it also known by discrete Karhunen Loève transforms (KLT), Hotelling transform or proper orthogonal decomposition (POD). In pattern recognition, computer vision, PCA is widely used for data representation and classical approach for feature extraction in various fields. Sirovich and Kirby (1987) first used PCA to efficiently represent pictures of human faces. First attempt of PCA for face recognition in the context of Eignefaces (1991) by the Turk and Pentland. They reconstructed the face image weighted sum of collection of all images and overall mean image. Therefore, PCA has been widely examined and has become one of the most successful approaches in face recognition.

3.3.3 KPCA

KPCA is a nonlinear generalized PCA, which performs on an arbitrarily large di-mension to select an appropriate feature space. There is no need to provide a number of features to select in advance like PCA, and it gives elucidation very clear when data is in high dimension [18]. Step by step explanation of KPCA algorithm is given bellow:

1. Get the Data in the form of a matrix
   $$\mathrm{Km} * \mathrm{m} \tag{4}$$
2. Performa Dot Product
   $$K_{ij} = (k(m_i, m_j))ij \tag{5}$$
3. Diagonal K and normalize the Eigenvector Expansion coefficients
   $$1 = \Lambda_n (\alpha^n . \alpha^n) \tag{6}$$
   Where α denotes the column vector with entries
   $\alpha 1 \ldots\ldots\ldots\ldots \alpha_n$.

As K is symmetric, it has a set of Eigenvectors which spans the whole space.

4. Compute Projection Matrix
   $$(\mathrm{kPC})_n = (Vn.\phi(\mathrm{m})) = \sum_{i=1}^{n} \propto_i^n k(\mathrm{m_i}, \mathrm{m}) \tag{7}$$

Where, ϕ is input face. The contour lines of constant projections onto the principal Eigenvector become nonlinear in input space

3.3.4 LDA

Linear Discriminant Analysis (LDA) is commonly used techniques for data classification and dimensionality reduction. Linear Discriminant Analysis easily handles the case where the within-class frequencies are unequal and their performances have been examined on randomly generated test data. This method maximizes the ratio of between-class variance to the within-class variance in any particular data set thereby guaranteeing maximal separability.

Linear discriminant analysis has been widely used both for data reduction and for enhancing the data separation ability as one of the most important subspace analysis method. This comes from the fact that not only within-class scatter but also between-class scatter are taken into account, leading to a highly effective solution to many pattern classification problems.

The term “discriminant” here refers to the discrimination among conditional distributions, whereas “association” refers to the two random vectors. A variable that discriminates well the conditional distributions captures association between the random elements. The prototype for this approach with two jointly Gaussian vectors is—as mentioned before—CCA with the canonical variables as linear discriminant functions. CCA is to be recovered as a special case. The term “linear discriminant function” was used by Kullback (1959/1968, p. 196f) to describe functions that are efficient in discriminating between two multivariate Gaussian distributions [24, 24].

Step 1: Formulate the data sets and the test sets which are to be classified in the original space. The given data sets and the test vectors are formulated. For ease of understanding let us represent the data sets as a matrix consisting of features in the form given below:

$$\textbf{Set 1=}\begin{bmatrix} \boldsymbol{a11} & \cdots & \boldsymbol{a12} \\ \vdots & \ddots & \vdots \\ \boldsymbol{am1} & \cdots & \boldsymbol{am2} \end{bmatrix} \qquad \textbf{(8)}$$

$$\textbf{Set 2=}\begin{bmatrix} \boldsymbol{b11} & \cdots & \boldsymbol{b12} \\ \vdots & \ddots & \vdots \\ \boldsymbol{bm1} & \cdots & \boldsymbol{bm2} \end{bmatrix} \qquad \textbf{(9)}$$

Step 2: Compute the mean of each data set and mean of entire data set.

Compute the mean of each data set and mean of entire data set. Let $\mu1$ and $\mu2$ be the mean of set 1 and set 2 respectively and $\mu3$ be mean of entire data, which is obtained by merging set 1 and set 2, is given by Equation 8.

$$\mu3 = \mu1\text{Xp1} + \text{p1 X}\mu2 \qquad \textbf{(10)}$$

Where and are the apriori probabilities of the classes. In the case of this simple two class problem, the probability factor is assumed to be 0.5.

Step 3: Calculate Within Class & Between Class

In LDA, within-class and between-class scatter are used to formulate criteria for class separability. Within-class scatter is the expected covariance of each of the classes. The Scatter measures are computed using Equations 11 and 12.

$$\boldsymbol{S_w} = \textstyle\sum_j \boldsymbol{p_j}\ \textbf{X}\ (\boldsymbol{Cov_j}) \qquad (11)$$

Therefore, for the two-class problem,

$$S_w = 0.5\ \text{X}\ Cov_1 + 0.5\ \text{X}\ Cov_2 \qquad (12)$$

All the covariance matrices are symmetric. Let and be the covariance of set 1 and set 2 respectively. Covariance matrix is computed using the following equation.

$$\boldsymbol{Cov_j} = (\boldsymbol{x_j} - \mu_j)(\boldsymbol{x_j} - \mu_j)^{\boldsymbol{T}} \qquad (13)$$

The between-class scatter is computes using the following equation

$$\boldsymbol{S_b} = \textstyle\sum_j (\mu_j - \mu_3)\textbf{X}(\mu_j - \mu_3)^{\boldsymbol{T}} \qquad (14)$$

Note that $\boldsymbol{S_b}$ can be thought of as the covariance of data set whose members are the mean vectors of each class. As defined earlier, the optimizing criterion in LDA is the ratio of between-class scatter to the within-class scatter. The solution obtained by maximizing this criterion defines the axes of the transformed space. However for the class-dependent transform the optimizing criterion is computed using equations (13) and (14). It should be noted that if the LDA is a class dependent type, for L-class separate optimizing criterion are required for each class. The optimizing factors in case of class dependent type are computed as

$$\boldsymbol{Criterion_j} = \boldsymbol{inv}(\boldsymbol{Cov_j})\ \textbf{X}\ \boldsymbol{S_b} \qquad (15)$$

For the class independent transform, the optimizing criterion is computed as

$$\boldsymbol{Criterion} = \boldsymbol{inv}(\boldsymbol{S_w})\ \textbf{X}\ \boldsymbol{S_b} \qquad (16)$$

Step 4: By definition, an eigen vector of a transformation represents a 1-D invariant subspace of the vector space in which the transformation is applied. A set of these eigen vectors whose corresponding eigen values are non-zero are all linearly independent and are invariant under the transformation. Thus any vector space can be represented in terms of linear combinations of the eigen vectors. A linear dependency between features is indicated by a zero eigen value. To obtain a non-redundant set of features all eigen vectors corresponding to non-zero eigen values only are considered and the ones corresponding to zero eigen values are neglected. In the case of LDA, the transformations are found as the eigen vector matrix of the different criteria defined in Equations (15) and (16). For any L-class problem we would always have L-1 non-zero eigen values. This is attributed to the constraints on the mean vectors of the classes in Equation 9. The eigen vectors corresponding to non-zero eigen values for the definition of the transformation. Once the transformations are completed using the LDA transforms, Euclidean distance or RMS distance is used to classify data points. Euclidean distance is computed using Equation 16 where $\mu_{\boldsymbol{ntrans}}$ is the mean of the transformed data set, is the class index and is the test vector. Thus for classes, Euclidean distances are obtained for each test point.

$$\boldsymbol{dist_n} = \left(\boldsymbol{transform_{n_{spec}}}\right)^{\boldsymbol{T}} \textbf{X}\ \boldsymbol{x} - \mu_{\boldsymbol{ntrans}} \qquad (17)$$

Step 5: The smallest Euclidean distance among the n distances classifies the test vector as belonging to class n.

3.3.5 Gabor Wavelet Transform

The characteristics of the Gabor wavelets, particularly the representations of frequency and orientation are suitable for representation and discrimination of texture information on the basis of human visual system and have been proven [20].

A Gabor filter can essentially be viewed as a complex exponential modulated by a Gaussian function i.e. a two-dimensional Gabor wavelet is a Gaussian kernel function modulated by a sinusoidal plane wave and can be represented as follows [21].

$$\Psi\omega,\Theta\ (\text{x},\text{y}) = \frac{1}{2\Pi\sigma 2\exp\left(-\left(\text{x}'2+\frac{\text{y}'2}{2\sigma 2}\right)\right)\exp(\text{jwx}')} \qquad (8)$$

Where $\text{x}' = \text{x}\cos\Theta + \text{y}\sin\Theta$ and $\text{y}' = -\text{x}\sin\Theta + \text{y}\cos\Theta$

where (x, y) is the position of the pixel in the spatial domain, the orientation of Gabor filter is denoted by Θ, the radial center frequency of the sinusoidal plane wave is denoted by ω and σ represents the standard deviation of the round Gaussian function(sharpness) along the co-ordinate axes. An ideal value for the standard deviation is suggested to be: σ≈π/ω and is

capable of distinctly defining the relationship between σ and ω.
Choosing 5 frequencies (m = 1,2,..5) and 8 orientations (n = 1,2,…8) for the Gabor filter bank results in the following equations:

$$\omega_m = \left(\frac{\Pi}{2}\right) x\sqrt{} - (m-1) \quad (9)$$

$$\theta_n = \left(\frac{\pi}{8}\right) x(n-1) \quad (10)$$

The change in the phase of the Gabor feature representation i.e. G $_{m,n}$ (x,y) is linear with small displacement of the sinusoid direction, but its magnitude change with respect to displacement is slow therefore the magnitude of the convolution outputs is assured.
The Gabor feature representation of an image I(x,y) is essentially the convolution of the image with the Gabor filter bank $\psi\ (x, y, \omega_m, \theta_n\ )$ and can be represented as follows:

$$0_{mn}(x,y) = I(x,y) \otimes \psi(x,y\ \bar{w}_{m,}\Theta_n) \quad (11)$$

3.4 Classification

Similarity measurement is playing an important role because it calculates the rank of the each similarity score and on the basis of that it calculates the rank level of each sub class. Therefore Mahalanobis Cosine (MAHCOS) is used for similarity measure [22]. It is also called as Cosine Mahalanobis Distance. The similarity measure between the testing image and the gallery image can be defined as the Cosine Mahalanobis Distance between the projections of the query and the gallery images. The effectiveness of MAHCOS has been demonstrated in [22].

Here m-by-n data matrix X, which is treated as m (1-by-n) row vectors x1, x2, ..., x_m, the various distances between the vector x_r and x_s are as follows:

$$d_{rs}^2 = (x_r - x_s)v^{-1}(x_r - x_s)^T$$

Classification on the basis of Similarity Measures here d.

3.5 Decision

Acceptance and rejection of the probe image is based on the feature extracted from the particular algorithm then calculate the similarity score of each individual and decision is taken on the basis of pre defined threshold.

4. EXPERIMENTAL RESULTS

4.1 Database

Biometrics Face Database is collection of human face images captured by cameras with certain parameters. ORL database is freely available on the web.

4.1.1 ORL Database

Source: This database is constructed by AT&T Laboratories Cambridge.
Purpose: This database is primarily used for face recognition.

Table 1 ORL Database

Properties	Descriptions
# of subjects	40
# of images	400
Static/Videos	Static
Single/ Multiple	Single
Gray/Color	Eight-bit gray
Resolution	92*112
Face pose	Moderate pose variation (up and down, quarter-profile to frontal-view)
Facial expression	3 facial expressions: neutral, smiling, closed eye
Illumination	N/A
Accessories	Glasses
3D data	N/A
Ground truth	Cropped face region Identifications of subjects

4.1.2 KVKR-Face Database

Source: This database is constructed by KVKR-Face Database under UGC SAP (II) DRS Phase-I & Phase-II, in the Multimodal Research Laboratory, Department of Computer Science and Information Technology, Dr. Babasaheb Ambedkar Marathwada University, Aurangabad. In this database two sensors are used to capture the face images i.e. IBall & Intex, separately. Detail description is given bellow in Table 2 is applicable for each sensors respectively.
Purpose: Biometrics Multimodal.

4.2 Approach

We have used KFA, Gabor+KFA, KPCA, Gabor+PCA, Gabor+KPCA, Gabor +LDA techniques for feature extraction details are explained in the section 3. Detail comparative results are shown in the Table 3 to 5.

Table 2 KVKR-Face Database

Properties	Descriptions
# of subjects	60
# of images/videos	600
Static/Videos	Both
Single/Multiple faces	Multiple
Gray/Color	Color
Resolution	640*480
Face pose	Normal, looking left, looking right, looking up, looking down with 45^0
Facial expression	Neutral, small smile, big smile, closed eye
Illumination	N/A
Accessories	Glasses, beards, moustaches
3D data	N/A
Ground truth	N/A

Table 3 Experiment 1(iBall Sensor)

Algorithm	rank one recognition rate	equal error rate	verification rate at 1% FAR	verification rate at 0.1% FAR
KFA	90.71%	5.95%	75.95%	50.71%
Gabor + KFA	35.56%	19.99%	37.22%	28.89%
GABOR PCA	92.50%	0.00%	100.00%	100.00%
GABOR KPCA	82.22%	0.04%	100.00%	100.00%
GABOR LDA	99.17%	0.04%	100.00%	100.00%
KPCA	76.67%	2.15%	97.14%	95.00%

Table 4 Experiment 2(Intex Sensor)

Algorithm	rank one recognition rate	equal error rate	verification rate at 1% FAR	verification rate at 0.1% FAR
KFA	94.52%	4.05%	89.29%	74.76%
Gabor + KFA	88.33%	2.22%	93.33%	81.11%

GABOR PCA	85.00%	0.83%	99.17%	96.67%
GABOR KPCA	74.44%	0.56%	99.44%	95.56%
GABOR LDA	97.50%	0.31%	100.00%	97.50%
KPCA	64.52%	2.96%	94.52%	84.05%

Performance via the cumulative matching characteristic (CMC) curve. The CMC curve represents the expectation of finding accurate match in the top r matches. In other words, a rank-r recognition rate shows the percentage of the test images that are correctly recognized from the top r matches in the gallery set. The rank-1 value on this curve indicates the true identification performance, while the rank- N score (N being the number of images in the gallery) will be 100% for closed-set re-identification, with the curve monotonically increasing from 1 through N, this approach is reported in the CMC curve in Fig.4.

Table 5 Experiment 3(ORL Database)

Algorithm	rank one recognition rate	equal error rate	verification rate at 1% FAR	verification rate at 0.1% FAR
KFA	85.71%	7.22%	81.07%	60.71%
Gabor + KFA	93.33	2.51%	96.67%	93.33%
GABOR PCA	74.17%	2.50%	97.50%	92.50%
GABOR KPCA	80.00%	4.17%	95.00%	92.50%
GABOR LDA	93.33%	2.51%	97.50%	93.33%
KPCA	49.29	9.29	69.29%	51.43%

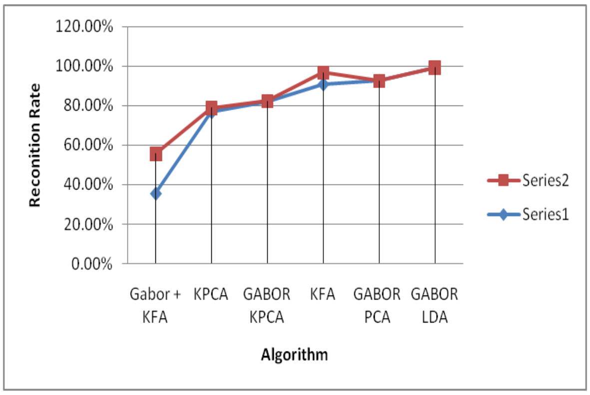

Fig.2 Comparative Result of iBall Sensor

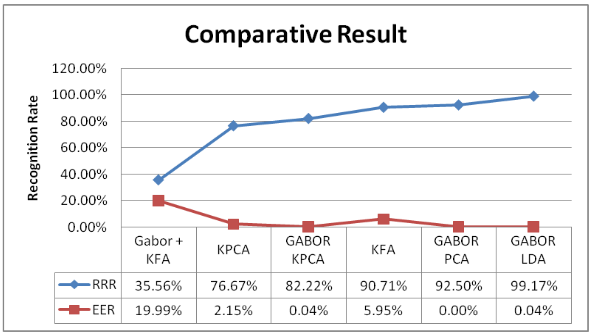

Fig.3 Comparative Result of Intex Sensor

As per the reported in the [23] Gabor +KFA gives 78% in FRGC we got improved recognition rate i.e. 88.33% on KVKR- face database and on ORL 93.33%.

4.3 Contributions

The Biometric database collection is a continuous and time consuming process where the database collectors have to work under different data acquisition frameworks and follow international standards. We have contributed a novel and standard database for face recognition i.e. KVKR- Face Database to the biometrics research community. Again, we have evaluated the face recognition systems on KVKR- and ORL face database. This includes a novel multi-algorithmic approach like Gabor+KPCA. The proposed systems on different databases have performed best and given significant results. This is also a contribution towards the biometric research community. The detail results are concluded in the following section.

5. CONCLUSIONS

In this paper used various algorithms and achieved better recognition rate. Results are compared of with sensor to sensor and it is proved that results are varying when use the different acquisition sensors. As compare to previous work cited in the literature it is proven and trusted techniques in various environment as well as indifferent sensors. Gabor + LDA give 99.17% RR. Compare results of ORL database with KVKR- Face database and gain enhanced RR. In future the technique should be implemented which gives constant result on various environments as well as on different sensors.

6. ACKNOWLEDGMENTS

The Author would like to acknowledge and thanks to the Programme Coordinator for providing me all necessary facilities and access to Multimodal Biometrics System Development Laboratory to complete this work under UGC SAP (II) DRS Phase-I & II F. No.-3-42/2009 & No.F.4-15/2015/DRS-II(SAP-II), One Time Research Grant No.F.19-132/2014 (BSR). We would like to acknowledge and thank to University Grants Commission (UGC), New Delhi for awarding me “UGC Rajiv Gandhi & BSR & MANF National Fellowship”.

7. REFERENCES

[1] K. Jain, Anil and Ross, Arun A and Nandakumar, *Introduction to biometrics*, no. September. Springer Science & Business Media, 2011.

[2] A. Ross, K. Nandakumar, and A. Jain, “Introduction to multibiometrics,” *Handb. Biometrics*, pp. 271–292, 2008.

[3] M. Grgic and K. Delac, “Face Recognition Home.” [Online]. Available: http://face-rec.org/databases/. [Accessed: 09-Aug-2016].

[4] C. Herrmann, “Extending a local matching face recognition approach to low-resolution video,” *2013 10th IEEE Int. Conf. Adv. Video Signal Based Surveillance, AVSS 2013*, no. 13, pp. 460–465, 2013.

[5] S. Prasad, T. Priya, M. Cui, and S. Shah, “Person Re-identification with Hyperspectral Multi-Camera Systems --- A Pilot Study,” 2016.

[6] S. J. Raudys and A. K. Jain, “Small sample size effects in statistical pattern recognition: recommendations for practitioners,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 13, no. 3. pp. 252–264, 1991.

[7] A. Jain and D. Zongker, “Feature Selection: Evaluation, Application, and Small Sample Performance,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 19, no. 2, pp. 153–158, 1997.

[8] J. Wang, K. N. Plataniotis, J. Lu, and A. N. Venetsanopoulos, “Kernel quadratic discriminant analysis for small sample size problem,” *Pattern Recognit.*, vol. 41, no. 5, pp. 1545–1555, 2008.

[9] K. Chougdali, M. Jedra, and N. Zahid, “Kernel relevance weighted discriminant analysis for face recognition,” *Pattern Anal. Appl.*, vol. 13, no. 2, pp. 213–221, 2010.

[10] L. Qiao, S. Chen, and X. Tan, “Sparsity preserving discriminant analysis for single training image face recognition,” *Pattern Recognit. Lett.*, vol. 31, no. 5, pp. 422–429, 2010.

[11] J. Wang, K. N. Plataniotis, J. Lu, and A. N. Venetsanopoulos, “On solving the face recognition problem with one training sample per subject,” *Pattern Recognit.*, vol. 39, no. 9, pp. 1746–1762, 2006.

[12] M. Ko and A. Barkana, “A new solution to one sample problem in face recognition using FLDA,” *Appl. Math. Comput.*, vol. 217, no. 24, pp. 10368–10376, 2011.

[13] S. Chen, J. Liu, and Z. H. Zhou, “Making FLDA applicable to face recognition with one sample per person,” *Pattern Recognit.*, vol. 37, no. 7, pp. 1553–1555, 2004.

[14] H. Yin, P. Fu, and S. Meng, “Sampled FLDA for face recognition with single training image per person,” *Neurocomputing*, vol. 69, no. 16–18, pp. 2443–2445, 2006.

[15] Q. xue Gao, L. Zhang, and D. Zhang, “Face recognition using FLDA with single training image per person,” *Appl. Math. Comput.*, vol. 205, no. 2, pp. 726–734, 2008.

[16] J. R.-S. J. Ruiz-del-Solar and P. N. P. Navarrete, “Eigenspace-based face recognition: a comparative study of different approaches,” *IEEE Trans. Syst. Man, Cybern. Part C (Applications Rev.*, vol. 35, no. 3, pp. 315–325, 2005.

[17] H. Jia and A. M. Martinez, “Face recognition with occlusions in the training and testing sets,” *2008 8th IEEE Int. Conf. Autom. Face Gesture Recognition, FG 2008*, 2008.

[18] X. Tan, S. Chen, Z. H. Zhou, and F. Zhang, “Face recognition from a single image per person: A survey,” *Pattern Recognit.*, vol. 39, no. 9, pp. 1725–1745, 2006.

[19] C. M. Coelho, S. Cloete, and G. Wallis, “The face-in-the-crowd effect: when angry faces are just cross(es).,” *J. Vis.*, vol. 10, no. 1, pp. 7.1–14, 2010.

[20] A. Vinay, V. S. Shekhar, K. N. B. Murthy, and S. Natarajan, “Face Recognition Using Gabor Wavelet Features with PCA and KPCA - A Comparative Study,” *Procedia Comput. Sci.*, vol. 57, pp. 650–659, 2015.

[21] A. Vinay, V. S. Shekhar, K. N. B. Murthy, and S. Natarajan, “Performance Study of LDA and KFA for Gabor Based Face Recognition System,” *Procedia Comput. Sci.*, vol. 57, pp. 960–969, 2015.

[22] Z. Fan, M. Ni, M. Sheng, Z. Wu, and B. Xu, “Principal Component Analysis Integrating Mahalanobis Distance for Face Recognition,” *2013 Second Int. Conf. Robot. Vis. Signal Process.*, pp. 89–92, 2013.

[23] A. Serrano, I. M. de Diego, C. Conde, and E. Cabello, “Recent advances in face biometrics with Gabor wavelets: A review,” *Pattern Recognit. Lett.*, vol. 31, no. 5, pp. 372–381, 2010.

[24] Daniela M. Witten, Robert Tibshirani, “Penalized classification using Fisher’s linear discriminant”, *Journal of the Royal Statistical Society, Series B Statistical Methodology*, 73, Part 5, pp. 753–772,2011

[25] Angelika van der Linde, “Dimension Reduction with Linear Discriminant Functions Based on an Odds Ratio Parameterization”, *International Statistical Review*, 71, 3, Pg. No. 629–666, 2003

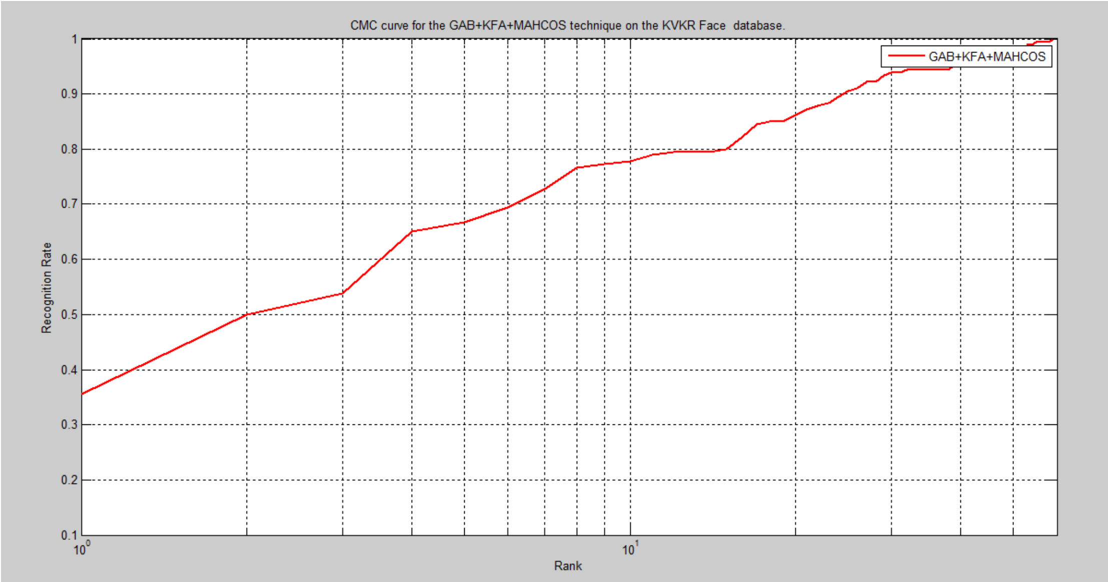

Fig.4 CMC Curve for the GAB+KFA+MAHACOS Technique on the KVKR- Database

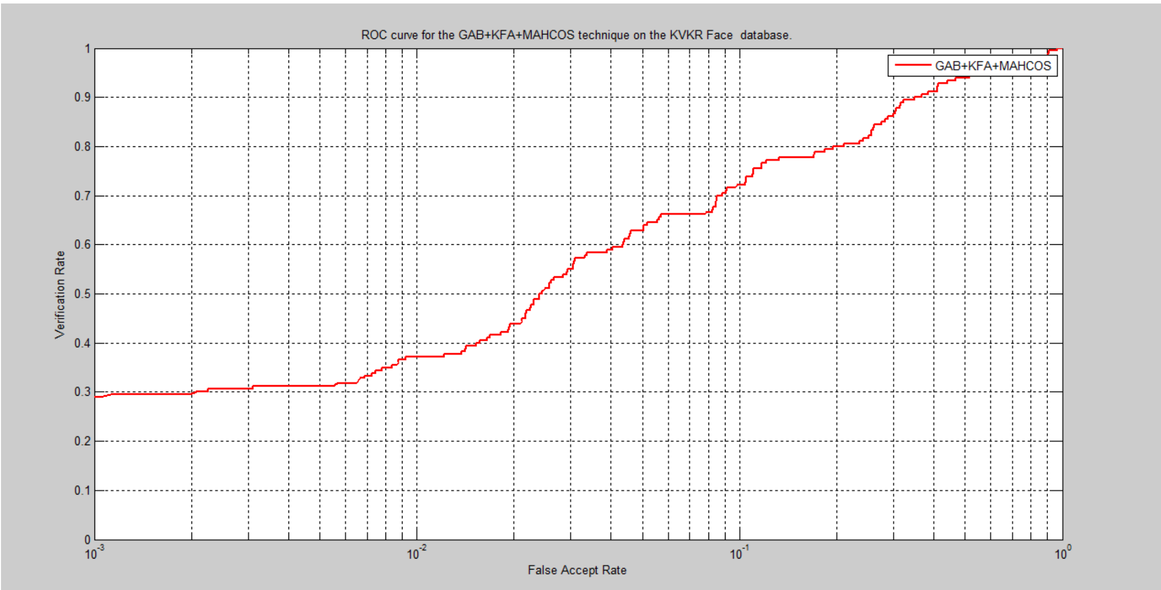

Fig.5 ROC Curve for the GAB+LDA+MAHACOS Technique on the KVKR- Database using iBall Sensor