International Journal on Emerging Technologies **11**(3): 681-685(2020)

ISSN No. (Print): 0975-8364
ISSN No. (Online): 2249-3255

Reliable Face Recognition System using KPCA based Feature vector

S. Princy Suganthi Bai[1], D. Ponmary Pushpa Latha[2], D. Joseph Pushpa Raj[3] and R. David Vinodhkumar Paul[4]

[1]Research Scholar, Bharathiar University, Coimbatore, Tamilnadu, India.
[1]Assistant Professor, Department of Computer Science, Sarah Tucker College, Tirunelveli (Tamilnadu), India.
[2]Associate Professor, Department of Information Technology, Karunya Institute of Technology and Sciences, Coimbatore (Tamilnadu), India.
[3]Assistant Professor, Department of Information Technology, Francis Xavier Engineering College, Tirunelveli (Tamilnadu), India.
[4]Assistant Professor, Department of Computer Science, St. John's College, Tirunelveli (Tamilnadu), India.

(Corresponding author: D. Ponmary Pushpa Latha)
(Received 27 March 2020, Revised 07 May 2020, Accepted 09 May 2020)
(Published by Research Trend, Website: www.researchtrend.net)

ABSTRACT: In the recent era, biometric-based security systems are topping in the monitoring and surveillance tasks in most of the commercial and public places. The face is easily readable and verifiable biometric which has undergone a decade of research. The main research scope in designing an FRS (Face Recognition System) is to bring out the best feature and classifier that produces high accuracy. The existing feature extraction methodologies LBP, PCA, DCT, FFT, DWT, etc., extract more number of feature vectors that increases the number of computations and focus on the fewer number of challenges. In this paper, a single UIEV (Unique Identifier Eigen Value) feature vector is proposed that is unique for each subject of the face database which is extracted through KPCA (Kernel Principal Component Analysis) methodology and classified with 1-NN rule (1-Nearest Neighbour) classifier. The novel approach of the UIEV feature vector is tested on five benchmark face databases to prove its reliability against the challenges like pose, accessories, lighting conditions, expression and time variations. The performance of the proposed approach reaches the optimal accuracy on all the face databases.

Keywords: Accuracy, Biometrics, Eigen Value, Face Recognition System, Kernel Principal Component Analysis.

Abbreviations: FFT, Fast Fourier Transform; DCT, Discrete Cosine Transform; PCA, Principal Component Analysis; LBP, Local Binary Pattern; GW, Gabor Wavelets; NN, Neural Networks; SVM, Support Vector Machine; KNN, K-Nearest Neighbours; LRC, Linear Regression Classification; LDA, Linear Discriminant Analysis; UIEV, Unique Identifier Eigen Value; 2DPCA, Two Dimensional PCA; CSGF, Circular Symmetrical Gabor Filter;

I. INTRODUCTION

Face recognition using a machine-learning approach is a remarkable development in the field of biometrics. Feature extraction and classification are the two important stages in the Face Recognition System (FRS). Many methodologies are proposed by the researchers to improve the FRS to overcome challenges like pose, expressions, noise, occlusion [1], age [5], illumination [8], time and lighting conditions [9].

The information from the face images is acquired globally or locally from the face images in the FRS. The acquired information is manipulated to get the important features in the feature extraction phase. Fast Fourier Transform (FFT) [10], Discrete Cosine Transform (DCT) [11, 12] and Principal Component Analysis (PCA) [1, 2, 4] are some of the Global based feature extraction methods in FRS models. Local Binary Pattern (LBP) [9, 14] and Gabor Wavelets (GW) [9] are the feature extraction methods used in FRS models that focus on the local features of the face images.

Classifiers utilized in FRS are also given equal weightage similar to the feature extraction methods. The role of classifiers in FRS is vital that supports the FRS to produce accurate results. The commonly used classifiers in FRS are Neural Networks (NN) [13, 26], Support Vector Machine (SVM) [1], K-Nearest Neighbours (K-NN) [1, 3, 6, 7, 9], Linear Regression Classification (LRC) [8] and Linear Discriminant Analysis (LDA) [10,12].

Generally, researchers test their FRS models with the benchmark face bases such as ORL, Yale, JAFFE, Indian, FERET, CMU PIE, AR, etc., that incorporates various challenges on face images. Also working with the standard face databases is worthwhile since the performance of the proposed FRS models can be compared and analysed reasonably with the existing FRS models.

Usage of the PCA technique for feature extraction in FRS stands top by producing good results than feature extraction methodologies such as LBP, FFT, DCT, GW, etc. The principal components of the face images are retrieved as features by using the PCA technique with different approaches to improving FRS efficiency in recent research works. Traditional PCA [15] converts the digital 2D face image in matrix form into 1D-column vectors, which occupies high dimension space. Yang *et al.,* (2004) in their work, proposed a two Dimensional PCA (2DPCA) novel approach in which the features are extracted as 2D matrix form to overcome the shortage of high dimensionality in traditional PCA [16]. This new approach of using 2DPCA is also in need of more coefficients that leads to space complexity and data redundancy. So the functionality of PCA on face images are further extended its function by working simultaneously on both the row and column directions and termed as (2D) 2PCA [17].

Circular Symmetrical Gabor Filter CSGF (2D)2PCA Neural Networks approach proposed by Kong *et al*., (2018) [1] is used in the feature extraction phase which undergoes deep learning on face images and brings out the high-level features that give a remarkable recognition in the training phase with SVM classifier. The novel approach of deep learning networks overcomes the issues in FRS such as data redundancy, computation time and no rotation invariance. The standard databases ORL, AR, Extend Yale B and XM2VTS are utilized to test their proposed FRS models that incorporate challenges like occlusion, illumination, pose, noise, expression variation, etc. The results of their FRS models produce good recognition rates of 97.50% with the ORL face database, 98.58% with AR face database, 99.58% with XM2VTS face database and 100% with Yale B face database that proves its sturdiness against these challenges but require more computations.

Zhou and Zhang (2018) constructed an FRS model based on the 2DPCA approach by reducing the sin angle between the input vectors and its variance [3]. This 2DPCAapproach reduces the input coefficients and produced a normal result even with noises. The proposed model reaches the maximum accuracy of 67.57% with the Extended Yale B database and 84.58% accuracy with the CMU PIE database. These results are not encouraging since other existing methodologies BA2DPCA [1], PCA [11] and LBP [14] reached the accuracy above 90%.

Abuzneid and Mahmood (2018) in their new FRS model incorporated LBP with histogram and multiple KNN classifiers for feature extraction [14]. The five different distance metrics such as Correlation, Euclidean, Canberra, Manhattan, and Mahalanobis are used to form a novel T-Dataset for classification. The Neural Network classifier with Back Propagation algorithm (BPNN) is utilized for classification which produces FRS accuracies of 93% with ORL face database and 91.1% with the Yale face database. Even though the performance of LBP with histogram and multiple KNN classifiers feature extraction technique is quite appreciable but the extended process to acquire T-Dataset is more time-consuming.

In this paper, the proposed FRS models used Kernel-based PCA to extract a single feature in the feature extraction phase, which requires less space for storage and unique for each subject so it is termed as a Unique Identifier Eigen Value (UIEV). The existing improved FRS models with PCA based feature extraction methodologies are lacking in storage aspect because of using more number of principal components. The 1-NN (Single Nearest Neighbour) rule classifier which is used for classifying the extracted features in the proposed FRS model is also more efficient by means of producing minimum computations in the classification process. The novel strategy of the UIEV feature extraction process produces a dominant feature for designing a robust FRS model that works well with ORL [18], Yale [19], JAFFE [21], NIR [22, 23] and Indian [24] face databases.

The usage of a 1-NN classifier in the proposed work follows a template matching approach that avoids regular training and testing phases which minimize the classification time. The arrangement of UIEV and 1-NN in designing FRS models needs less space and time complexity. The proposed strategy to design a flawless FRS model is described in section II. The experiments conducted with the proposed methodology on the suggested standard face databases are illustrated in section III and section IV includes the concluding part of this paper. The future scope of the proposed work is added in section V.

II. PROPOSED METHODS

In the proposed FRS model, KPCA[20] approach is used in the feature extraction phase to extract the UIEV feature vector from each face image of the face databases, which are unique for a particular subject. The extracted features are then classified with a 1-NN rule classifier that matches the test template with the first nearest matching template in the training space. The first nearest neighbour (k=1) rule considers the first closest matching record available in the feature space. The 1-NN rule avoids template matching with next level neighbours of k values such as two or three or more that will pull down the accuracy. A polynomial kernel with the PCA approach is the best strategy to extract the UIEV feature vector that needs less space and time complexity in the training phase and produces good accuracy in the testing phase. In this section, the KPCA procedure to extract the UIEV feature vector and matching technique of 1-NN rule are well defined.

Traditional KPCA Procedure: A face image is a matrix grid of pixel values. Each face image x_i is transformed from the original dimensional space D to the high dimension space M as $\varphi(x_i)$. The transformed values $\varphi(x_i)$ are summed and divided by the total number of samples N. Then the transformed data mean value is set as zero that is expressed in the Eqn. (1).

$$\frac{1}{N}\sum_{i=1}^{N}\varphi(x_i)=0 \tag{1}$$

The statistical metric covariance C^F is calculated with the use of the transformed values $\varphi(x_i)$ and transposed values T of transformed values $\varphi(x_i)^T$ in the non-linear feature space(F) which is shown in the Eqn. (2).

$$C^F=\frac{1}{N}\sum_{i=1}^{N}\varphi(x_i)^T\varphi(x_i) \tag{2}$$

Eigen values(λ) and Eigen vectors(v_k) are manipulated with the covariance C^F in the M dimension space that is given in the Eqn. (3) where k=1,2,..M.

$$C^F v_k=\lambda v_k \tag{3}$$

The Eigen vector $v_k \epsilon F$ is expressed with coefficients a_k in Eqn. (4) where λ is not equal to zero.

$$v_k=\sum_{i=1}^{N}a_{ki}\ \varphi(x_i) \tag{4}$$

The kernel principal components $y_k(x)$ can be calculated by applying polynomial kernel with d degree which is shown in the Eqns. (5) and (6).

$$k(x,y)=(x^Ty)^d \tag{5}$$

$$y_k(x)=\sum_{i=1}^{N}a_{ki}\ k(x,x_i) \tag{6}$$

The collected components are processed to elevate the UIEV feature vector from the face images that is elaborated in the next sub-division.

Extraction of UIEV feature vector process: Initially, traditional KPCA model steps are utilized to get the sorted Eigen values in the extraction of UIEV feature vector process. The KPCA technique which performs dimension reduction involves several steps to extract Eigen vectors and Eigen values in FRS. The Eigen values obtained for each face image by the KPCA approach are sorted and the first three Eigen values are isolated for each sample face image. The best and nearest Eigen value for all sample images of the same subject is stored as the feature vector for that face image. This process is done by calculating the similarity of Eigen values among the same class with Euclidean distance metric in 1-NN rule methodology. The mean value of the eigenvalues of the previous subject is calculated and used as the reference value while fixing the eigenvalue of the next new subject to have maximum inter-subject distance. Thus the eigenvalues of each subject are unique and called a UIEV feature vector. The feature extraction process of the proposed FRS model is shown in Fig. 1.

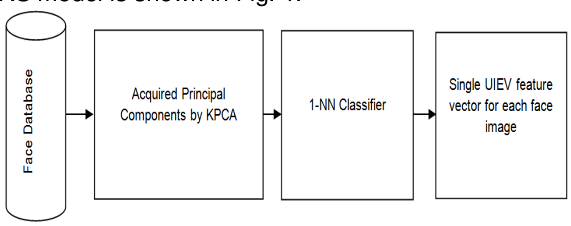

Fig. 1. The feature extraction process of the proposed FRS model.

Matching technique of 1-NN rule: Nearest-neighbour classifier [25] is a distance-based method, which is applied to large datasets for the application like pattern recognition. In this approach, the training tuples are placed in the feature space excluding the testing tuple. Then the matching is performed by measuring the distance between the training tuples and testing tuple. Generally, K holds the number of matches considered for analysing, which varies from one to any numeric value. Probably in FRS, K varies from one to three [14]. In the proposed work, the first best closest tuple (k=1) is considered for the performance evaluation and Euclidean distance is used for the distance calculation given in Eqn. (7).

$$\mathrm{d}(\mathrm{T}_1, \mathrm{T}_\mathrm{i}) = \sqrt{(t_1 - t_i)^2} \tag{7}$$

Here in the above equation, t_1 is the UIEV feature of testing tuple T_1, t_i is the UIEV features of all other training tuples T_i and i indicates each training sample in the feature space.

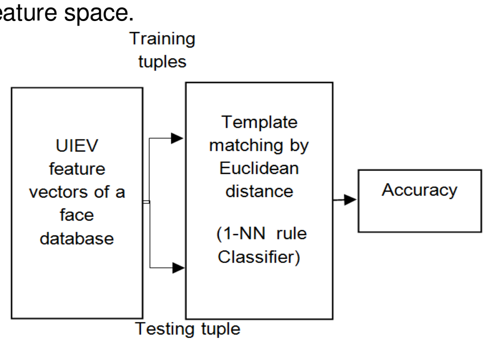

Fig. 2. Classification phase of the proposed FRS model.

The classification phase of the FRS procedure produces the accuracy of the FRS model which is depicted in Fig. 2.

The performance of the proposed FRS models is evaluated with the accuracy metric calculated by Eqn. (8) in the classification phase.

$$\text{Accuracy} = \frac{\text{Total number of correct matches}}{\text{Total number of samples}} \times 100 \tag{8}$$

III. RESULTS AND DISCUSSION

The proposed FRS model of using the UIAE feature vector is classified with a 2-KNN and 1-KNN rule classifier. The experiments are conducted on five face databases which include the challenges like scale, pose, illumination variation, expression, light conditions, etc.

The information about the five databases used in the experiments is summarized in the Table 1. The UIEV features are extracted from the face databases using MATLAB17. The accuracy is measured with a 2-KNN and 1-KNN rule classifier in the IBM SPSS modeler.

Table 1: Summary of five face databases.

Face database	Number of subjects	Challenges incorporated
Yale	15	Accessories, diverse lighting conditions and expressions.
NIR	15	Expression, pose and time variations
ORL	40	Accessories, lighting conditions, expressions and time.
JAFFE	10	Emotional face expressions and uniqueness in appearance
Indian	59	Emotional facial expressions and poses.
Yale	15	Accessories, diverse lighting conditions and expressions.

In the classification phase, the cross-validation [25] approach is used, in which the training and testing samples are allocated and deal located in the various folds. This approach divides the dataset into different subsets and one subset is dedicated for testing, remaining other samples are used for training. All the subsets will be provided with the chance of being in testing and training space by shuffling the subsets between the training and testing space. Template matching is the technique performed by the NN rule classifier in the classification phase. The UIAE features extracted from the five face databases are experimented with the newly designed FRS models with 2-NN rule classifier and 1-NN rule classifier.

In the proposed work, each dataset of five face databases is distributed into 5-folds that reduce the number of computations while comparing with 10-folds. This concept makes every subset occur in a training subset and testing subset that raises the accuracy to 100% with a 1-NN classifier for all the face databases chosen for the experiments.

The experiments prove that the single feature vector UIEV with 1-NN rule stands robust to any type of challenge that exists in the face images. Even though many different distance metrics are available, the Euclidian distance metric is used because of its promising results in the FRS recognition process.
The outcomes of the FRS models with a 2-NN rule classifier show a good response range from 94% to 97%,except for the Yale face database. In this case, while considering the first nearest neighbour alone in the KNN classifier, all the five face databases reached the optimal accuracy.
The performance of the FRS models experimented on the five face databases with diverse K values as one and two are recorded carefully in Table 2 and the comparison is shown in Fig. 3.

Table 2: Accuracy comparison of five face databases with diverse K values.

Face database	Accuracy with KNN-5 folds K=2	Accuracy with KNN-5 folds K=1
Yale	49.70%	100%
NIR	94.67%	100%
ORL	96%	100%
JAFFE	96.50%	100%
Indian	96.90%	100%

The proposed FRS model when compared with the recent approaches gains good attraction because of its performance and advantages over other existing FRS models that are described in Table 3 and Fig. 4.

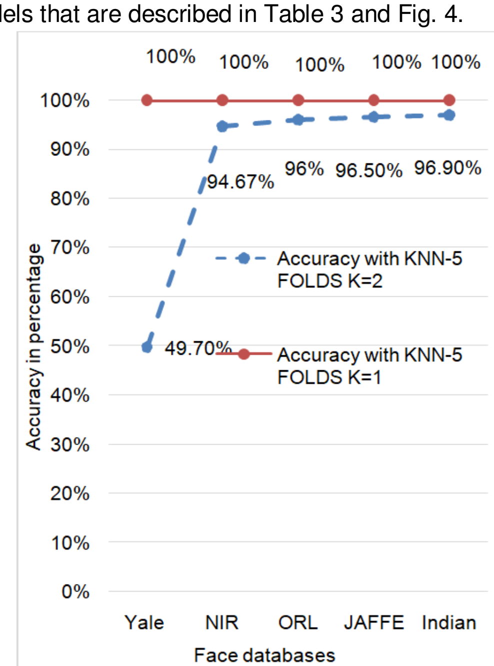

Fig. 3. Accuracy comparison on five face databases with diverse K values.

The usage of LBP and traditional PCA in the feature extraction phase of designing the FRS models show good accuracy, but not able to reach the optimal accuracy. Generally, the researchers experimented with their FRS models [1, 3, 11, 14] with the fewer face databases and scored less reasonable accuracy that ranges between 82.95% to 97.5%.

The optimal result has been achieved with this proposed Eigen based FRS model. It is noted that UIEV vector and 1-NN classifier can effectively handle any challenges of FRS. The poor performance of the FRS model with the Yale face database also resolved with this proposed model.

Table 3: Comparison of Proposed model with recent FRS models.

S. No.	Author	Technique and face database	Accuracy
1.	Cheng-Yaw Low *et al.*, (2017) [7]	PCA + Hebbian Learing + ORL	97.0%
2.	Abuzneid & Mahmood (2018) [14]	LBP + KNN + Yale	91.1%
3.	Abuzneid & Mahmood (2018) [14]	LBP + KNN + ORL	93%
4.	Kong *et al.*, (2018) [1]	CSGF(2D)2 + PCA + SVM + ORL	97.5%
5.	Zhou & Zhang (2019) [3]	BA2DPCA + 1-NN + Yale (with noise)	82.95%
6.	Proposed model (2019)	UIEV + 1-NN + five face databases	100%

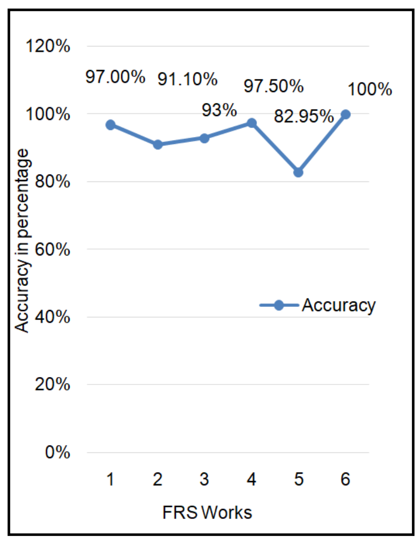

Fig. 4. Accuracy Comparison of Proposed model with recent FRS models.

IV. CONCLUSION

In this paper, the proposed UIEV feature vector and 1-KNN classifier are incorporated to design an FRS model. The 1- NN classifier serves in two ways in this proposed FRS model. It acts as a feature vector selector in the feature extraction phase by extracting a single best KPCA based UIEV feature vector and as a classifier in the classification process. The reliability of the proposed model is assessed with five different databases with diverse challenges. The response of this novel approach is found to be robust irrespective of challenges like pose, time, illumination variation and expression variation when the classifier value K is set as one.

V. FUTURE SCOPE

People are wearing masks in public and commercial places due to the threat of corona virus and other medical reasons so that the whole face image is not revealed for the recognition process in the FRS. This occlusion challenge must also be considered while designing the FRS which is incorporated with the AR face database. The outcomes of the proposed FRS model gives a good scope to test its performance in the other face databases with occlusion challenges because of the necessity to solve the real-world FRS problems.

ACKNOWLEDGEMENTS

We record our sincere thanks to our affiliated institutions for their motivations and suggestions to complete the work in time.

REFERENCES

[1]. Kong, J., Chen, M., Jiang, M., Sun, J., & Hou, J. (2018). Face recognition based on CSGF (2D) 2 PCANet. *IEEE Access*, *6*, 45153-45165.

[2]. Wang, Q. (2012). Kernel Principal Component Analysis and its Applications in Face Recognition and Active Shape Models.*ArXiv*.

[3]. Zhou, S., & Zhang, D. (2018). Bilateral Angle 2DPCA for Face Recognition. *IEEE Signal Processing Letters*, *26*(2), 317-321.

[4]. Shah, J. H., Sharif, M., Raza, M., & Azeem, A. (2013). A Survey: Linear and Nonlinear PCA Based Face Recognition Techniques. *Int. Arab J. Inf. Technol.*, *10*(6), 536-545.

[5]. Atallah, R. R., Kamsin, A., Ismail, M. A., Abdelrahman, S. A., & Zerdoumi, S. (2018). Face recognition and age estimation implications of changes in facial features: A critical review study. *IEEE Access*, *6*, 28290-28304.

[6]. Lai, Z., Xu, Y., Chen, Q., Yang, J., & Zhang, D. (2014). Multilinear sparse principal component analysis. *IEEE transactions on neural networks and learning systems*, *25*(10), 1942-1950.

[7]. Low, C. Y., Teoh, A. B. J., & Toh, K. A. (2017). Stacking PCANet+: An overly simplified convnets baseline for face recognition. *IEEE Signal Processing Letters*, *24*(11), 1581-1585.

[8]. Huang, S. M., & Yang, J. F. (2012). Improved principal component regression for face recognition under illumination variations. *IEEE signal processing letters*, *19*(4), 179-182.

[9]. Tan, X., & Triggs, B. (2010). Enhanced local texture feature sets for face recognition under difficult lighting conditions. *IEEE transactions on image processing*, *19*(6), 1635-1650.

[10]. Hwang, W., Wang, H., Kim, H. Seok-CheolKee, & Kim, J. (2011). Face Recognition System Using Multiple Face Model of Hybrid Fourier Feature Under Uncontrolled Illumination Variation. *Transactions on Image Processing, 20*(4), 1152-1165.

[11]. Atta, R., & Ghanbari, M. (2010). Low-memory requirement and efficient face recognition system based on DCT pyramid. *IEEE Transactions on Consumer Electronics*, *56*(3), 1542-1548.

[12]. Ezoji, M., & Faez, K. (2011). Use of matrix polar decomposition for illumination-tolerant face recognition in discrete cosine transform domain. *IET Image Processing*, *5*(1), 25-35.

[13]. Sudha, N., Mohan, A. R., & Meher, P. K. (2011). A self-configurable systolic architecture for face recognition system based on principal component neural network. *IEEE transactions on circuits and systems for video technology*, *21*(8), 1071-1084.

[14]. Abuzneid, M. A., & Mahmood, A. (2018). Enhanced human face recognition using LBPH descriptor, multi-KNN, and back-propagation neural network. *IEEE Access*, *6*, 20641-20651.

[15]. Jolliffe. T. (2002). Principal Component Analysis, *Springer-Verlag*, Second Edition.

[16]. Yang, J., Zhang, D., Frangi, A. F., & Yang, J. Y. (2004). Two-dimensional PCA: a new approach to appearance-based face representation and recognition. *IEEE transactions on pattern analysis and machine intelligence*, *26*(1), 131-137.

[17]. Zhang, D., & Zhou, Z. H. (2005). (2D) 2PCA: Two-directional two-dimensional PCA for efficient face representation and recognition. *Neurocomputing*, *69*(1-3), 224-231.

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

[19]. Yale face database: http:// cvc.cs.yale.edu/cvc/ projects/yalefaces/yalefaces.html

[20]. Wang, W., Zhang, M., Wang, D., & Jiang, Y. (2017). Kernel PCA feature extraction and the SVM classification algorithm for multiple-status, through-wall, human being detection. *EURASIP Journal on Wireless Communications and Networking*, (1), 1-7.

[21]. JAFFE Yale face database: http:// www.kasrl.org/ jaffe.html

[22]. Zhang, B., Zhang, L., Zhang, D., & Shen, L. (2010). Directional binary code with application to PolyU near-infrared face database. *Pattern Recognition Letters*, *31*(14), 2337-2344.

[23]. NIR face database:http:// www.comp. polyu.edu .hk/~biometrics/NIRFace/polyudb_face.htm

[24]. Indian face database: http:// viswww.cs.umass.edu /~vidit/ IndianFace database

[25]. Han, J., Pei, J., & Kamber, M. (2011). Data Mining Concepts and Techniques. Third Edition, *Elsevier*, 348–350.

[26]. Salmam, F. Z., Madani, A., & Kissi, M. (2018). Emotion recognition from facial expression based on fiducial points detection and using neural network. *International Journal of Electrical and Computer Engineering*, *8*(1), 52-59.

How to cite this article: Bai, S. P. S., Latha, D. P. P., Raj, D. J. P. and Paul, R. D. V. (2020). Reliable Face Recognition System using KPCA based Feature Vector. *International Journal on Emerging Technologies*, *11*(3): 681–685.