

Research Article

Investigation of Feature Selection Techniques for Face Recognition Using Feature Fusion Model

¹N. Vijaya Kumar and ²M.S. Irfan Ahmed

¹Research and Development Centre, Bharathiar University,

²Sri Krishna Institutions, Sri Krishna College of Engineering and Technology, Coimbatore, India

Abstract: This study investigates various feature selection techniques for face recognition. Biometric based authentication system protects access to resources and has gained importance, because of their reliable, invariant and discriminating features. An automated biometric system is based on physiological or behavioral human characteristics for protected access. Biometric trait such as palmprint, iris, hand, voice, face fingerprint, or signature is used to authenticate a person's claim. Of the biometrics, face recognition is gaining popularity due to its simple method of capturing the image using cameras. However the number of features generated is high leading to higher computation time. Using feature selection technique it is shown that recognition rate improves.

Keywords: Biometrics, Correlation-based Feature Selection (CFS), face recognition, k Nearest Neighbor (kNN), Mutual Information (MI), Naive Bayes (NB), ORL face database

INTRODUCTION

Biometrics identifies people by measuring their individual anatomy or physiology (hand geometry/fingerprint), deeply ingrained skill, behavioural characteristic (signature) or a combination of the two (like voice). Biometric authentication technologies like face, hand, finger, iris and speaker recognition are available and in use. A biometric system is a pattern recognition system that acquires an individual's biometric data extracting a feature set and comparing it with a database template (Podio, 2002).

Based on the context, a biometric system operates in verification or identification modes. In the verification mode, the system corroborates a person's identity by comparing captured biometric data with own biometric template (s) in the system database. In the identification mode, the system identifies an individual by searching the database for all users' templates for a match. The two types of biometrics are Unimodal and Multimodal Biometrics.

The unimodal rely on evidence of one information source for authentication (single fingerprint, face) (Ross and Jain, 2004; Anwar *et al.*, 2009). Some possible issues in this type of biometric are:

- Noise in sensed data
- Intra-class variations
- Inter-class similarities
- Non-universality
- Spoof attacks

The term multi modal biometrics denotes fusion of different information types (Anwar *et al.*, 2009) (e.g., fingerprint and face of same person or fingerprints from two diverse fingers of a person). Multi biometrics addressed issues related to unimodal like (Ross *et al.*, 2008):

- Non-universality or insufficient population coverage (reducing failure to enroll rate that increase population coverage).
- It is increasingly difficult for impostors to spoof multiple biometric traits of legitimately enrolled individuals.
- Multi biometric systems address noisy data issues (illness affecting voice; scar affecting fingerprint).

Multi biometric systems offer improvement in matching accuracy of biometric systems based on information being combined and fusion methodology adopted (Teoh *et al.*, 2004). Advances in Information Technology make Information Security inseparable. Authentication plays an important role in dealing with security. Information security is concerned with integrity, confidentiality and information availability in all forms. There are many tools and techniques supporting information security management. But systems based on biometrics evolved supporting some information security aspects. Biometrics with reference to biological sciences was studied and applied for generations and is now viewed as "biological statistics."

Corresponding Author: N. Vijaya Kumar, Research and Development Centre, Bharathiar University, Coimbatore, India

This work is licensed under a Creative Commons Attribution 4.0 International License (URL: <http://creativecommons.org/licenses/by/4.0/>).

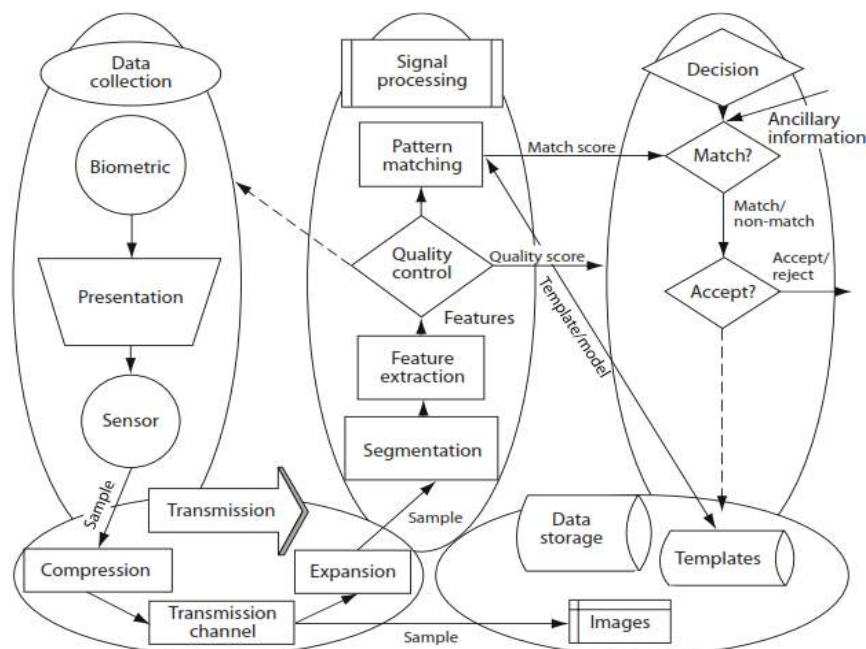


Fig. 1: A generic biometric system

Biometric authentication supports identification, authentication and non-repudiation in information security. Biometric authentication is popular in providing personal identification that is important in applications like credit card fraud and identity theft which indicate that this is an issue of major concern. Authentication confirms something (or someone) as true, i.e. that claims made by or about a thing is true (Council, 2007; Bhattacharyya *et al.*, 2009).

A complete biometric system is characterized by three elements namely.

Enrolment sub-system: Here, data samples are collected from an enrollee. Mostly devices like scanners and readers are used for this. This stage is crucial as mistakes lead to identity misrepresentation.

Template representation: At this biometric operation stage, data samples obtained during enrolment are gathered and stored for future reference. Gathering and Storing is usually done by specific software tools.

Matching process subsystem: Input data is compared with stored data templates in the system for identification and verification.

Figure 1 shows a generic biometric authentication system. It generally has five sub-systems: data collection, transmission, signal processing, decision and data storage. Biometric systems begin with measurement of a behavioural or physiological characteristic (data collection). Some biometric systems collect data at one location but store and process it at another. They need data transmission. When much data is involved, compression is needed before transmission or storage to save bandwidth and storage space. Signal-

processing subsystems are split into four tasks: feature extraction, segmentation, quality control and pattern matching. Features are extracted from the biometric image which is used for matching to authenticate. Segmentation step is required for some biometric like iris to select only the portion of the biometric image required for authentication. Quality control checks whether the image received from the sensor is of good quality and features can be extracted. In case of defective image, it requests for a new sample. The remaining subsystem is storage. There are one/many forms of storage used, based on the biometric system. Decision subsystem implements system policy by directing database search. It determines “matches” or “non-matches” based on distance or similarity measures from pattern matcher and makes an “accept/reject” decision based on system policy (Jain *et al.*, 2004).

Biometrics is classified into physiological biometrics and behavioural biometrics. Physical Characteristics such as Face, Fingerprint, Iris, Palmprint, Retina and behavioural Characteristics like Signature, Voice, Keystroke, Gait are commonly used biometrics (Kumar and Ryu, 2008). Different technologies are used for face recognition. One approach captures a face image using inexpensive camera (visible spectrum). This models key features from the central portion of a facial image extracting features from captured image (s) that do not change with time and avoids superficial features like facial expressions or hair. Major facial recognition benefits include being non-intrusive, hands-free, ensuring continuous authentication and acceptance by users. A face recognition system is a computer vision that automatically identifies a human face from database images. Face recognition problem is challenging as it

has to account for all possible appearance variation due to illumination, facial features and occlusions change. Face recognition is used for two tasks (Latha *et al.*, 2009):

- Verification (one-to-one matching)
- Identification (one-to-many matching)

Face images acquisition depends on underlying applications. For instance, surveillance applications capture face images by a video camera while image database investigations require static intensity images from a standard camera. Face recognition's advantage of ubiquity and being universal over other biometrics is that everyone's face is displayed readily. Uniqueness, another biometric characteristic is hard to claim at current accuracy levels. As face shape, specially when young, is influenced by genotype, identical twins are hard to tell apart with this technology.

Geometric feature based approaches are earliest approaches to face recognition and detection (Lu *et al.*, 2003). Here, significant facial features are detected and distances among them and other geometric characteristics are joined in a feature vector to represent the face. To recognize a face, test image feature vector and database image should be obtained. Template based approaches represent a technique to detect faces (Nefian and Hayes III, 1999). Unlike geometric feature based approaches template based approaches use feature vector representing the face template rather than significant facial features. Correlation based face detection methods are based on the computation of normalized cross correlation coefficient (Kekre *et al.*, 2011). The first step is determining the location of significant facial features like eyes, nose or mouth. Importance of robust facial feature detection for detection and recognition resulted in development of various facial feature detection algorithms. Philips introduced a template based face detection/recognition system using a matching pursuit filter to get face vector (Phillips, 1998). Matching pursuit algorithm applied to image iteratively selects from a basis functions dictionary the best image decomposition by reducing the image residue in iterations.

This study investigates the impact of feature fusion and feature selection for face recognition. Features are extracted using Gabor for texture feature and DCT energy coefficients. The features are fused and the best features selected using Correlation based Feature (CFS) selection and Mutual Information (MI).

LITERATURE REVIEW

Face recognition accuracy was improved by reducing uncertainty by Yong *et al.* (2014), to reduce face representation uncertainty by synthesizing virtual training samples. Useful training samples similar to test

sample from a set of original and synthesized virtual training samples are used. It devised a selected useful training samples based representation approach to perform face recognition. Results on five widely used face databases prove that this approach obtained high face recognition accuracy, with lower computational complexity than state-of-the-art approaches.

Developing a face recognition system based on PCA and Self-Organizing Maps (SOM) unsupervised learning algorithm was focused on by Anggraini (2014). Preprocessing contains grey scaling, cropping and binarization. Selected dataset for research was Essex database at the University of Essex consisting of 7900 face images from 395 individuals (male and female).

An up-to-date critical survey of still and video based face recognition research was provided by Zhao *et al.* (2003). There are two underlying motivations to write this survey paper: the first providing an up-to-date review of current literature and the next is to offer insights into studies of machine face recognition. To ensure a comprehensive survey, it categorized present recognition techniques and presented detailed descriptions of representative methods in every category. Also, relevant topics like psychophysical studies, system evaluation and illumination issues and variation pose are covered.

Machine learning methods for recognition based on face and iris biometrics was dealt with by Oravec (2014). It presented relevant machine learning methods with focus on Neural Networks (NN). Some NN theory aspects like visualization of processes in NN, input data internal representations as base for new feature extraction methods and classification and compression applications were addressed. Machine learning methods are used for feature extraction and classification and are applicable to biometric systems.

A multi-modal face recognition algorithm was presented by Jihua *et al.* (2013). After Feature extraction using 2DPCA and dimension reduction from multiple two-dimensional face images (a positive face and 2 side faces) of a person, it identified through a new feature Matrix reformed by part of a feature vector of every 2 dimensional face images. Result of CAS-PEAL Face Database revealed that this method had higher recognition rate than that which used one positive face under same conditions.

A local approach for face recognition based on combined feature selection methods like Genetic Algorithm (GA), mRmR features selection algorithm, GramdtShmidt algorithm and Naive Bayesian classifier was proposed by Ouarda *et al.* (2013) which was compared with global features based face recognition systems. This study gives a comparative study based on Recognition rates and Execution times. The Naive Bayesian classifier based face recognition system tested on ORL face database showed 78.75% recognition rate

and interesting execution times compared to global approaches.

Nazari and Moin (2013) proposed new face recognition based on fusing global and local features. To extract global and local features, it used Gabor wavelet filter to apply on entire image and non-overlapping sub-images with equal size. To reduce new fused feature vector dimension, PCA is used. In experiments, it used kNN and multi-class SVM classifiers and ORL database to get face recognition rate. Results reveal that the new face recognition algorithm outperforms conventional methods like global Gabor face recognition and G-2DFLD feature fusion face recognition regarding recognition rate.

An effective Multi-Sub region Correlation Filter Bank (MS-CFB) based feature extraction algorithm for robust face recognition proposed by Yan *et al.* (2014) combined benefits of global and local-based feature extraction algorithms where multiple correlation filters corresponding to different face sub regions are designed jointly to optimize overall correlation outputs. It reduced MS-CFB computational complexity by designing a correlation filter bank in spatial domain and improved generalization capitalizing on unconstrained form during filter bank design process.

A system that collects face images at large standoff in both daytime and nighttime and presents an Augmented Heterogeneous Face Recognition (AHFR) method for cross-distance (150 m probe vs. 1 m gallery) and cross-spectral (visible light gallery vs. near-infrared probe) face matching was reported by Kang *et al.* (2014). It recovered high-quality face images from degraded probe images by proposing a Locally Linear Embedding (LLE) based image restoration method based. Restored face images are matched to gallery using a heterogeneous face matcher. Results showed that new AHFR approach outperformed state-of-the-art methods for cross-spectral and cross-distance face matching.

A new noise modeling framework to improve representation based classification for robust face recognition was proposed by Zhang *et al.* (2015). Representation based classification evoked repercussions in the face recognition field. The new framework iteratively first diminishes representation noise achieving better representation solution for linear combination till it converges and then exploits determined 'optimal' representation solution and fusion method to perform classification. Experiments proved that the new framework simultaneously improves representation capability by decreasing representation noise and improving RCBM classification accuracy.

A sparse representation based classification proposed by Tang *et al.* (2014) also suggested a new classification method called Weighted Group Sparse Representation Classification (WGSRC) to classify query images by minimizing weighted mixed-norm (ℓ_2, ℓ_1 -norm), regularized reconstruction error regarding

training images. It represented a test sample by training samples from neighbors and highly relevant classes. WGSRC sparse solution encodes more structure information and discriminative information than other sparse representation methods. Results on five face data sets revealed that the new method outperformed state-of-the-art sparse representation based classification methods.

A new face recognition using Extended Curvature Gabor (ECG) Classifier Bunch described by Hwang *et al.* (2014) extended Gabor kernels into ECG kernels by adding a spatial curvature term to kernel and adjusting Gaussian width at kernel, leading to many feature candidates being extracted from one image. An ECG classifier is implemented by applying LDA to selected feature vector. To overcome accuracy limitation of one classifier, an ECG classifier bunch combining multiple ECG classifiers to fusion scheme is proposed.

METHODOLOGY

This study proposed face recognition using feature fusion. Features are extracted using Gabor for texture feature and DCT. Features are concatenated. Feature selection uses CFS, MI. Naive Bayes and k Nearest Neighbor (kNN) are classifiers. ORL face database is used for evaluating the recognition system.

ORL face database: AT and T (formerly ORL database) is a standard face database. AT and T database has face images of 40 distinct persons with each having 10 different images, taken at different times, totaling 400. Each database face image has 112×92 pixels size. Facial expressions have variations like open/closed eyes, smiling/non-smiling and facial details like glasses/no glasses. All images were taken against a dark homogeneous background with subjects in an up-right, frontal position, with tolerance for a few side movements. There are also scale variations. Though database was used in many face recognition researches, it is clear that many samples or database size are too small to prove eventual results. A higher size database is essential to prove the accuracy of face recognition researches.

Feature extraction:

Discrete Cosine Transform (DCT): Discrete Cosine Transform (DCT) (a Fourier-related transform), is real valued, which is implemented using the Discrete Fourier Transform (Rao *et al.*, 1990). The DCT computes a truncated Chebyshev series. It expresses the data in terms of sum of cosine functions. The common type of DCT used functions on a real sequence x_n of length N to produce coefficients C_k , as follows:

$$C_k = \frac{2}{N} w(k) \sum_{n=0}^{N-1} x_n \cos\left(\frac{2n+1}{2N} \pi k\right), \quad 0 \leq k \leq N-1$$

and,

$$x_n = \sum_{k=0}^{N-1} w(k) C_k \cos\left(\frac{2n+1}{2N} \pi k\right), \quad 0 \leq n \leq N-1$$

where,

$$w(k) = \sqrt{2}, \quad k=0 \text{ and } w(k) = 1, \quad 1 \leq k \leq N-1$$

The DCT has strong energy compaction property. DCT has been effectively used in face recognition (Hafed and Levine, 2001) instead of Karhunen-Loeve Transform (KLT) as DCT is computationally less intensive.

Gabor filters: Gabor filters are very useful in image processing, for it has optimal localization properties in both frequency and spatial domain (Daugman, 1985). The Gabor function is a harmonic oscillator present within a Gaussian envelope and composed of sinusoidal plane wave. A 2-D gabor filter over an image (x, y) can be defined as:

$$G(x, y) = \exp\left(-\frac{(x-x_0)^2}{2\sigma_x^2} - \frac{(y-y_0)^2}{2\sigma_y^2}\right) \times \exp(-2\pi i(u_0(x-x_0) + v_0(y-y_0)))$$

where,

- (x_0, y_0) = Specify location in image
- (u_0, v_0) = Specify modulation that has spatial frequency
- $\omega_0 = \sqrt{u_0^2 + v_0^2}$
- θ_0 = Orientation, $\theta_0 = \arctan(v_0/u_0)$
- σ_x, σ_y = Standard deviations

Gabor filters are used as bandpass filters to remove noise. To apply Gabor filters to an image, the frequency of the sinusoidal plane wave, the filter orientation and the standard deviations of the Gaussian envelope are to be specified.

Feature fusion: The principle of image fusion using wavelets is to merge the wavelet decompositions of the two original images using fusion methods like mean-mean, max-min, img1, img2 and mean-max applied to approximation coefficients and detail coefficients. Due to the difference in correlation among the features, the mean-max fusion method is used. The combined feature vectors reduced and classified using the Naive Bayes and KNN. Fusion at the feature level is used for selection and combination of features to eliminate redundant and irrelevant features.

Feature selection:

Correlation-based Feature Selection (CFS): CFS is a fast, correlation-based filter algorithm applied to continuous and discrete problems (Ross *et al.*, 2008).

CFS algorithm is a heuristic to evaluate merit of a features subset. This algorithm considers individual features usefulness to predict class label with inter-correlation level among them. The heuristic is based on the hypothesis: Good feature subsets have features correlated to class, yet uncorrelated to others. In test theory, the same principle designs a composite test (sum/average of individual tests) to predict an external variable of interest where features are individual tests measuring traits related to variable of interest (class):

$$Merit_s = \frac{\overline{kr_{cf}}}{\sqrt{k+k(k-1)r_{ff}}}$$

Mutual Information (MI): Entropy is a random variables uncertainty measure. Let X be a random variable with discrete values, its uncertainty are measured by entropy H (X), defined as (Peng *et al.*, 2005):

$$H(X) = -\sum_{x \in X} p(x) \log 2 p(x)$$

where, $p(x) = \Pr(X=x)$ is probability density function of X. Entropy does not depend on actual values, just the random variable's probability distribution. For two discrete random variables X and Y with their probability density function $p(x, y)$, joint entropy H (X; Y) is (Fleuret, 2004):

$$H(X, Y) = -\sum_{x \in X, y \in Y} p(x, y) \log 2 p(x, y)$$

Information shared between two random variables is mutual information. Given variable X, how much information one gains about variable Y, which is MI I (X; Y):

$$\begin{aligned} I(X; Y) &= H(X) - H(X|Y) \\ &= H(Y) - H(Y|X) \\ &= -\sum_{x \in X, y \in Y} p(x, y) \log 2 \frac{p(x, y)}{p(x)p(y)} \end{aligned}$$

According to above equation, the MI I (X; Y) will be big if two variables X and Y are related. If not, I (X; Y) = 0 if X and Y are unrelated (Estévez *et al.*, 2009).

Classifiers:

Naive Bayes: Naive Bayes classifiers are Bayes theorem based statistical classifiers using a probabilistic approach to predict data class, by matching it to class with highest posterior probability. The following algorithms are used in Naive Bayes:

$$P(C_i|V) = \frac{P(V|C_i)P(C_i)}{P(V)}$$

where, $V = (v_1, \dots, v_n)$ is document represented in n -dimensional attribute vector and C_1, \dots, C_m represents m class. But it is computationally expensive to compute $P(V|C_i)$ to reduce computation, so naïve assumption of class conditional independence is made. Hence:

$$P(V|C_i) = \prod_{k=1}^n P(x_k|C_i)$$

k-Nearest Neighbour (kNN): k-NN classifier's premise is that vector space model is similar for same documents. Training documents are indexed and associated with corresponding label. A test document when submitted, is treated as query retrieving documents from training set similar to test document. Test document class label assigning is based on k-NN distribution. Class label is refined by adding weights leading to higher accuracy by tuning. k-NN is simple to understand and easy to implement:

$$p(x) \cong \frac{k}{NV}$$

Similarly probability density function $p(x|H_i)$ of observation x is conditioned to hypothesis H_i can be approximated. It is assumed that N_i is number of patterns associated to hypothesis H_i , $i = 1 \dots C$, so that $N_1 + \dots + N_C = N$.

The example is classified by determining samples majority of labels for K-Near neighbor. This method is easy to enforce if an example "x" has k nearest examples where feature space and many of them have same label "y", then "x" belongs to "y". K-NN depends on further theorem when considering theory. Decision course considers small nearest neighbor number. So, example disproportion problem is solved when this method is used. Though limited nearest neighbor is considered by K-NN, it is not a decision boundary. Hence K-NN is suitable to classify example set of boundary intercross if it overlaps. Euclidian distance is calculated as follows. Two vectors x_i and x_j are given where $x_i = (x_i^1, x_i^2, x_i^3, x_i^4, x_i^5, \dots, x_i^n)$ and $x_j = (x_j^1, x_j^2, x_j^3, x_j^4, x_j^5, \dots, x_j^n)$.

The difference between x_i and x_j is:

$$D(x_i, x_j) = \sqrt{\sum_{k=1}^{n-1} ((x_{ik} - x_{jk}))^2}$$

RESULTS AND DISCUSSION

Thirty seven people's images are used for the conducting experiments to evaluate the proposed work. Ten images for each person were used and 37×5 images were used for training and equal amount for testing. The images are obtained from ORL database. Figure 2 shows the recognition rate when feature selection techniques are not used.

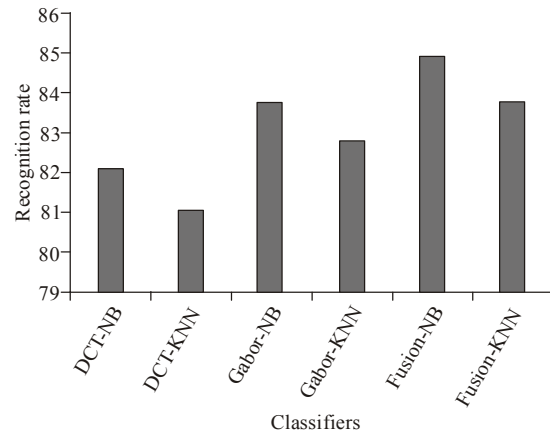


Fig. 2: Recognition rate when feature selection techniques are not used

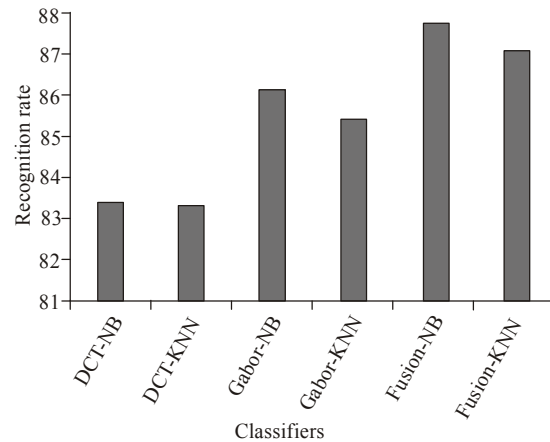


Fig. 3: Recognition rate when mutual information is used

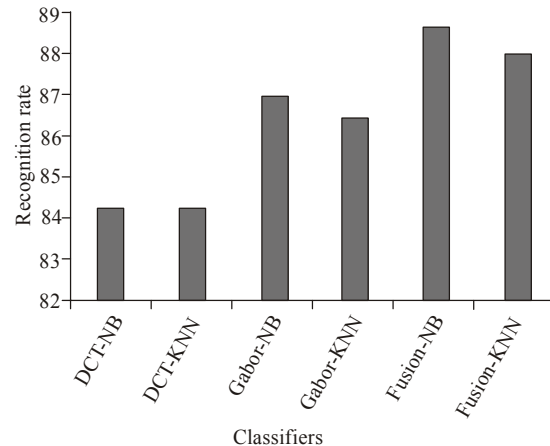


Fig. 4: Recognition rate when correlation-based feature selection is used

The proposed fusion technique with Naïve Bayes classifier approach improved the recognition rate by 4.56% when compared with DCT-NB method showing the effectiveness of the fusion technique. Similarly kNN improved the recognition rate by 3.28% when

compared with DCT-kNN method. Figure 3 shows the recognition rate when MI is used for feature selection.

It can be seen that MI improves the recognition rate compared to Fig. 2. The proposed fusion with Naïve Bayes approach improved the recognition rate by 3.81% when compared with DCT-NB method. The proposed fusion with kNN improved the recognition rate by 3.83% when compared with DCT-kNN method.

Figure 4 shows the recognition rate when CFS is used for feature selection. The performance is better than Mutual Information based feature selection though both are statistical techniques.

The proposed fusion with Naïve Bayes approach improved the recognition rate by 5.01% when compared with DCT-NB method. The proposed fusion with kNN improved the recognition rate by 4.4% when compared with DCT-kNN method.

CONCLUSION

Biometrics has the feature of exclusivity and unchanged, or acceptably changed, over life time of an individual that is deemed as one of the best solutions to control access. In this study, we used fusion for the face recognition. CFS and MI methods were used for the feature selection and the classifiers were kNN and Naïve Bayes. The experiments conducted in three scenarios like without feature section, with CFS and with MI for the DCT, Gabor and Fusion with the classifiers. The results demonstrated that the fusion is outperformed than the DCT and Gabor methods in all three scenarios.

REFERENCES

- Anggraini, D.R., 2014. Face recognition using principal component analysis and self organizing maps. Proceeding of the 3rd ICT International Student Project Conference (ICT-ISPC, 2014).
- Anwar, F., M.A. Rahman and S. Azad, 2009. Multibiometric systems based verification technique. Eur. J. Sci. Res., 34(2): 260-270.
- Bhattacharyya, D., R. Ranjan, F. Alisherov and M. Choi, 2009. Biometric authentication: A review. Int. J. u- e-Serv. Sci. Technol., 2(3): 13-28.
- Council, S.C.A.I., 2007. Identity and Smart Card Technology and Application Glossary. Smart Card Alliance Identity Council.
- Daugman, J.G., 1985. Uncertainty relation for resolution in space, spatial frequency and orientation optimized by two-dimensional visual cortical filters. JOSA A, 2(7): 1160-1169.
- Estévez, P.A., M. Tesmer, C.A. Perez and J.M. Zurada, 2009. Normalized mutual information feature selection. IEEE T. Neural Networ., 20(2): 189-201.
- Fleuret, F., 2004. Fast binary feature selection with conditional mutual information. J. Mach. Learn. Res., 5: 1531-1555.
- Hafed, Z.M. and M.D. Levine, 2001. Face recognition using the discrete cosine transform. Int. J. Comput. Vision, 43(3): 167-188.
- Hwang, W., X. Huang, S.Z. Li and J. Kim, 2014. Face recognition using extended curvature gabor classifier bunch. Pattern Recogn., 48(4): 1243-1256.
- Jain, A.K., A. Ross and S. Prabhakar, 2004. An introduction to biometric recognition. IEEE T. Circ. Syst. Vid., 14(1): 4-20.
- Jihua, Y., X. Guomiao and H. Dan, 2013. Research and implementation of of multi-modal face recognition algorithm. Proceeding of the IEEE 25th Chinese Control and Decision Conference (CCDC, 2013), pp: 2086-2090.
- Kang, D., H. Han, A.K. Jain and S.W. Lee, 2014. Nighttime face recognition at large standoff: Cross-distance and cross-spectral matching. Pattern Recogn., 47(12): 3750-3766.
- Kekre, H.B., S.D. Thepade, S. Khandelwal, K. Dhamejani and A. Azmi, 2011. Face recognition using multilevel block truncation coding. Int. J. Comput. Appl., 36(11).
- Kumar, D. and Y. Ryu, 2008. A brief introduction of biometrics and fingerprint payment technology. Proceeding of the 2nd International Conference on Future Generation Communication and Networking Symposia (FGCNS'08), 3: 185-192.
- Latha, P., L. Ganesan and S. Annadurai, 2009. Face recognition using neural networks. Signal Process. Int. J., 3(5): 153-160.
- Lu, J., K.N. Plataniotis and A.N. Venetsanopoulos, 2003. Face recognition using kernel direct discriminant analysis algorithms. IEEE T. Neural Networ., 14(1): 117-126.
- Nazari, S. and M.S. Moin, 2013. Face recognition using global and local Gabor features. Proceeding of the IEEE 21st Iranian Conference on Electrical Engineering (ICEE, 2013), pp: 1-4.
- Nefian, A. and M.H. Hayes III, 1999. A hidden Markov model-based approach for face detection and recognition. Ph.D. Thesis, School of Electrical and Computer Engineering, Georgia Institute of Technology.
- Oravec, M., 2014. Feature extraction and classification by machine learning methods for biometric recognition of face and iris. Proceeding of the IEEE 56th International Symposium on ELMAR (ELMAR), pp: 1-4.
- Ouarda, W., H. Trichili, A.M. Alimi and B. Solaiman, 2013. Combined local features selection for face recognition based on Naïve Bayesian classification. Proceeding of the IEEE 13th International Conference on Hybrid Intelligent Systems (HIS), pp: 240-245.

- Peng, H., F. Long and C. Ding, 2005. Feature selection based on mutual information criteria of max-dependency, max-relevance and min-redundancy. *IEEE T. Pattern Anal.*, 27(8): 1226-1238.
- Phillips, P.J., 1998. Matching pursuit filters applied to face identification. *IEEE T. Image Process.*, 7(8): 1150-1164.
- Podio, F.L., 2002. Personal authentication through biometric technologies. *Proceeding of the IEEE 4th International Workshop on Networked Appliances*. Gaithersburg, pp: 57-66.
- Rao, K.R., P. Yip and K.R. Rao, 1990. *Discrete Cosine Transform: Algorithms, Advantages, Applications*. Vol. 226, Academic Press, Boston.
- Ross, A. and A. Jain, 2004. Multimodal biometrics: An overview. *Proceeding of the 12th European Signal Processing Conference (EUSIPCO)*. Vienna, Austria, pp: 1221-1224.
- Ross, A., K. Nandakumar and A.K. Jain, 2008. Introduction to Multibiometrics. In: *Handbook of Biometrics*. Springer, US, pp: 271-292.
- Tang, X., G. Feng and J. Cai, 2014. Weighted group sparse representation for Undersampled Face Recognition. *Neurocomputing*, 145: 402-415.
- Teoh, A., S.A. Samad and A. Hussain, 2004. Nearest neighbourhood classifiers in a bimodal biometric verification system fusion decision scheme. *J. Res. Pract. Inf. Tech.*, 36(1): 47-62.
- Yan, Y., H. Wang and D. Suter, 2014. Multi-subregion based correlation filter bank for robust face recognition. *Pattern Recogn.*, 47(11): 3487-3501.
- Yong, X., F. Xiaozhao, L. Xuelong, Y. Jiang, J. You, L. Hong and T. Shaohua, 2014. Data uncertainty in face recognition. *IEEE T. Cybern.*, 44(10): 1950-1961.
- Zhang, Z., L. Wang, Q. Zhu, Z. Liu and Y. Chen, 2015. Noise modeling and representation based classification methods for face recognition. *Neurocomputing*, 148: 420-429.
- Zhao, W., R. Chellappa, P.J. Phillips and A. Rosenfeld, 2003. Face recognition: A literature survey. *ACM Comput. Surv.*, 35(4): 399-458.