

Face Recognition based on a Hybrid Meta-Heuristic Feature Selection Algorithm

S.Arivalagan

Department of Computer Science & Engineering,
Annamalai University,
Annamalainagar–608002, India.

K.Venkatachalapathy

Department of Computer Science & Engineering,
Annamalai University,
Annamalainagar–608002, India.

ABSTRACT

For the past few years, a number of new face recognition techniques have been proposed. Always it is a big unanswered question among face recognition researchers about which method or technique will have better performance. In this study an approach to recognize known faces based on Eigen vectors and a hybrid Meta-heuristic feature selection algorithm is proposed. The eigenvectors which are covariance matrix of the face images together describes the difference between face images. Face recognition problem is viewed as a two dimensional recognition problem. Initially the face images are projected in to face space and using Principal component analysis the eigenvectors with high Eigenvalues are extracted to reduce the dimension of the feature vector. Further to select the best feature vectors which increase the classification accuracy is selected by using a hybrid meta-heuristic algorithm using Genetic algorithm (GA) and Bacteria Foraging Optimization (BFO). In this study the Support vector machine (SVM) and Back propagation neural network (BPNN) are used for classification. The classifiers are trained and tested separately using the frontal face images taken from AT&T database. The SVM and BPNN produces an average classification accuracy of 82.6% and 83.28% respectively.

Keywords

Meta-Heuristics, Genetic Algorithm, Bacteria Foraging Optimization, support vector machine (SVM)

1. INTRODUCTION

Automatic face recognition by computers is now becoming popular area for research and there now many successful techniques are invented due to advancements in the field of computing over past few decades. Face recognition system involves finding mathematical representation of the face images and matching processes. A face recognition system can be used both verification and identification purposes. An extensive number of researches have been done already on the field of face recognition. This paper further explores the face recognition using Eigenfaces combined with a hybrid meta-heuristic approach for selecting the features. The known face images are classified using classifiers namely, Support Vector Machine (SVM) and Back Propagation Neural Network (BPNN) separately. The performance of the both the classifiers used are compared finally to identify the suitable classifier which could be used efficiently for face recognition problem. The sample frontal face images for training and testing SVM and BPNN are taken from the AT&T database. The database contains ten different images of each of 40 distinct subjects. All the images were taken against a dark homogeneous background with the subjects in an upright frontal position. In this study for training and testing the

classifiers 5 different images of each of 20 distinct subjects were taken from the AT&T database.

In the proposed facial recognition system initially the sample face images are transformed in to Eigen faces using principal component analysis (PCA) algorithm. There are patterns which can be observed in the domain of facial recognition due to the presence of objects like eyes, nose and mouth in face. These characteristics features are called Eigenfaces in the facial recognition domain. These features are extracted from the image data by using a mathematical tool called Principal Component Analysis [1]. The main idea of using PCA for face recognition is to express the large 1-D vector of pixels constructed from 2-D facial image in to the compact principal components of the feature space. This can be called projection of Eigenspace [2].

A hybrid Meta-Heuristic feature selection algorithm is used to find the optimal feature subset which improves the classification accuracy. In this paper a feature extraction technique using PCA and a feature selection technique using hybrid Meta-heuristics algorithm are integrated together to form a robust and efficient method to alleviate the effects of curse of dimensionality.

The rest of the paper is organized as follows. The previous and works related to face recognition are described in Section 2. In Section 3 the Eigenface computation using PCA technique is explained. The feature selection using hybrid meta-heuristics algorithm is explained in the Section 4. The Section 5 contains details about the two classifiers used in this study. The section 6 describes about the various experiments conducted on the images taken from AT&T database and the results obtained. Finally the section 7 concludes the work done in the above sections.

2. RELATED WORKS

Many of the previous work in the field of Face recognition depend majorly on the human facial features such as the eyes, mouth, nose and face edges. A combination of four different individual face recognition techniques was used by Zakariya S.M., et al.,[3] to develop a face recognition systems based on one combination of four individual techniques namely Principal Component Analysis (PCA), Discrete Cosine Transform (DCT), Template Matching using Correlation (Corr) and Partitioned Iterative Function System (PIFS).

Scale Invariant Feature Transform (SIFT) is always considered as a powerful technique for general object recognition/detection. In [4] Cong Geng, et al., proposed two new approaches Volume-SIFT (VSIFT) and Partial-Descriptor-SIFT (PDSIFT) for face recognition based on the original SIFT algorithm.

A hybrid method for face recognition using local features and statistical feature extraction methods is proposed by Donghyun kim, et al., in [5]. In that hybrid approach first, a dense set of local feature points are extracted in order to represent a facial image. Each local feature point is described by the keypoint descriptor defined by SIFT feature. Then, the statistical feature extraction methods, PCA and LDA, are applied to the set of local feature descriptors in order to find low dimensional features. With the obtained low dimensional feature vectors face recognition task is conducted efficiently using a simple classifier.

Many commercial vendors of facial recognition system claim the successive working of their system but the Face Recognition Vendor Test (FRVT) conducted by the National Institute of Standards and Technology (NISI), USA indicates that the commercial face recognition systems do not perform up to the mark under the variations ubiquitously present in a real-life situation [6]. An extensive number of researches have been carried globally to develop Facial Recognition system with high accuracy.

To overcome the problems caused due to high dimensionality of the feature set, many progressive research works have been done to produce optimal feature selection strategies.

The Feature Selection (FS) aims at finding the optimal set of d features out of m [7], [8] and [9]. Many methods have been proposed previously to perform feature selection on training and testing data. Among the various methods proposed for FS, population-based optimization algorithms such as Genetic Algorithm (GA)-based method [10], [11], [12] and Ant Colony Optimization (ACO)-based method have been utilized widely [13].

3. BACKGROUND CONCEPTS

3.1 Genetic Algorithm

Genetic algorithms (GAs) are widely used for variety of random and local search methods and they are categorized as adaptive search techniques. Over the years GA which is considered to be an inductive searching technique demonstrated substantial improvement [14]. The Feature selection using genetic algorithms are considered to be an excellent choice for improving the performance of the classification system [15]. The selection of an appropriate representation and an adequate evaluation function is vital for the successful implementation of any searching problem using GA. For the problem of feature selection the importance of selecting the appropriate representation and evaluation function is described in [16]. The binary representation is the most widely used and easy to implement method which represents each feature in the candidate feature set as a binary gene and each binary representation is a binary string of fixed length representing some subset of the given feature set. An initial feature set with irrelevant, redundant and noisy features are provided to the GA. This high dimensional feature set containing positive and negative examples of the various classes serves both as input and as well as a training set. Using a search procedure the space of all subsets of the given feature set is explored. Using the evaluation function the performance of each of the selected feature subsets is measured.

3.2 Bacteria Foraging Algorithm

Bacteria Foraging Optimization Algorithm (BFOA), proposed by Passino [18] is a novel optimization algorithm based on the

social foraging behavior of *E. coli* bacteria. The original Bacterial Foraging Optimization system consists of three principal mechanisms, namely, chemo taxis, reproduction, and elimination-dispersal [19]. In the **Chemotaxis** process the movement of an *E.coli* cell through swimming and tumbling via flagellais simulated.

In the **Reproduction process** the least healthy bacteria eventually die while each of the healthier bacteria are asexually split into two bacteria, which are then placed in the same location. This is been done to keep the swarm size constant.

In the **Elimination and Dispersal** process a gradual or sudden changes in the local environment may occur due to various reasons e.g. a significant local rise of temperature may kill a group of bacteria that are currently in a region with a high concentration of the nutrient gradients. Due to the occurrence of these kinds of events all the bacteria in a region are killed or a group is dispersed into a new location. To simulate this phenomenon in BFOA some bacteria are liquidated at random with a very small probability while the new replacements are randomly initialized over the search space [18].

4. EIGENFACE COMPUTATION USING PCA

In this section the steps involved in the Eigenface computation using PCA technique is briefly presented. The steps include preparing the data, subtracting the mean, calculating the covariance matrix, calculating the eigen vectors and eigen values of the covariance matrix and finally selecting the principal components. The sample faces for training and testing are taken from the AT&T database (formerly 'The ORL Database of Faces') for further processing. Totally 100 of 10 different subjects (10 of each) images are taken for processing. The size of each image is 92x112 pixels, with 256 grey levels per pixel.

Let the training set images be $\Gamma_1, \Gamma_2, \dots, \Gamma_M$. The average image has to be ψ has to be calculated and subtracted from each face image. The mean subtracted image is represented by ϕ .

$$\Psi = \frac{1}{M} \sum_{n=1}^N \Gamma_n \quad (i)$$

$$\phi_i = \Gamma_i - \Psi \quad (ii)$$

The calculate the co-variance matrix using formula

$$C = \frac{1}{M} \sum_{n=1}^M \phi_n \phi_n^T \quad (iii)$$

(or) $C = A.A^T$ where $A = [\phi_1, \phi_2, \dots, \phi_M]$.

Finally find the Eigenvalues and Eigenvectors of the covariance matrix. If the size of the face image is $M \times N$ then the size of the eigenvector will be $[M \times N] \times 1$. If the number of sample images taken for training and testing is increased then the size of the feature set will be high and the training time will also be higher. To avoid this eigenvectors with high Eigenvalues (i.e) the principal components are selected. The eigenvectors with high Eigenvalue best describes the

characteristics of a individual face than the eigenvectors with less eigenvalue.

5. FEATURE SELECTION USING HYBRID META-HEURISTIC ALGORITHM

The feature selection deals with the task of selecting the best feature set which reduces the classifier training time and as well as increase the classification accuracy. From a given a set of d features, the feature selection algorithm selects a subset of size m which increases the classification accuracy [20]. The Feature selection algorithm also aims at dimensionality reduction of data, which initially contain a high number of features. Thus the Feature selection allows the reduction of feature space, which is crucial in reducing the training time and improving the prediction accuracy [21]. This is achieved by removing irrelevant, redundant and noisy features from the initial feature set (i.e) it selects the subset of features that can achieve the best performance in terms of accuracy and computation time. It performs the dimensionality reduction. Most widely used feature selection algorithm uses various search procedures to select the best set of features. A number of search procedures have been proposed. Popularly used feature selection algorithms are Sequential forward Selection, Sequential Backward selection, Genetic Algorithm and Particle Swarm Optimization. In this work a hybrid approach using Genetic Algorithm combined with Information theory and Bacteria Foraging optimization is proposed to select the optimal features. The optimal features are then considered for classification.

5.1 Feature selection using Genetic Algorithms

If the size of the feature set is m then there will be 2^m possible feature subsets. The selection of best feature subset can be viewed as a combinatorial optimization problem and is solved using Genetic Algorithms as proposed in [22]. Each face image is transformed in to a vector of eigen-features which are the coefficients of the linear expansion of the image in the eigenspace. In the representation method used in this study, the chromosome is represented using a bit string whose length is same as the number of eigenvectors. For each eigenvector one bit in the string is associated. If the i^{th} bit is 1, then the i^{th} eigenvector is selected, otherwise, that component is ignored. Thus each chromosome represents a different eigen-feature subset. The main aim of using the feature selection algorithm is to use fewer features to attain the same or better classification accuracy. Therefore, the fitness evaluation contains two terms: (i) accuracy and (ii) number of features used. The fitness function is calculated as proposed in [23]

$$\text{fitness} = 10^4 \text{Accuracy} + 0.4 \times \text{Zeros}$$

where Accuracy is the classification accuracy rate that an individual feature set achieves, and Zeros is the number of zeros in the chromosome. If the accuracy is higher, then the fitness will also be higher.

5.2 Feature Selection using Bacteria Foraging Optimization

The BFO based feature selection method used in this study was proposed by Rasleen Jakhar, et al., which reduces the number of features, removes irrelevant, noisy and redundant data, and results in acceptable recognition accuracy [21]. In this paper the Eigenface of frontal face image are computed

using PCA technique. The resulting feature subset (obtained by BFO) is the most representative subset and is used to recognize the face from face gallery.

The steps involved in Feature selection using BFO algorithm is explained below:

Define the BFO parameters: S , N_c , N_{re} , n , where S : Number of bacteria in the colony, N_c : chemo tactic steps, N_{re} : reproductive steps, n : dimension of the search space

1. Initially place all the bacteria randomly at a position.
2. For each reproduction step perform the following operations
3. For each chemo taxis step perform the following operations

3.1 Compute the fitness function using the following equation

Let w_1, w_2, \dots, w_L and N_1, N_2, \dots, N_L denotes the classes and number of images within each class, respectively. Let M_1, M_2, M_L and M_0 be the means of corresponding classes and the grand mean in the feature space.

M_i can be calculated as:

$$M_i = \frac{1}{N_i} \sum_{j=1}^{N_i} W_j^{(i)}, \quad i = 1, 2, \dots, L$$

where $W_j^{(i)}$, $j=1, 2, \dots, N_i$, represents the sample image from class W_i .

The grand mean M_0 is computed using the formula below,

$$M_0 = \frac{1}{N} \sum_{i=1}^L N_i M_i, \quad \text{where } N \text{ is the}$$

total number of images of all the classes.

Thus the between class scatter fitness function F is computed as follows:

$$F = \sqrt{\sum_{i=1}^L (M_i - M_0)^t (M_i - M_0)}$$

3.2 Next change the bacteria to new position

3.3 Compute the fitness function for the new position as in step 3.1.

3.4 If fitness function of new position is less than the fitness of the previous position then move bacteria back to

its previous position. Else update the fitness function of the bacteria to the new position.

(chemo taxis loop of step 3 ends).

4. (reproduction step) Calculate the *health* of each bacteria and Sort the bacteria in descending order of *health*.

5. The bacteria with lowest health values are left and other bacteria with best *health* values are split and the newly generated bacteria copies are placed at the same location as their parents. (reproduction loop of step 4 ends)

6. Pick up the position of bacteria B with $\max(\text{health})$ value. This position represents the best feature subset of the features.

5.3 Proposed Hybrid Algorithm

Get the extracted features $f_1, f_2, f_3, \dots, f_N$ from Image I (x, y) and form the feature set as $F = \{ f_1, f_2, f_3, \dots, f_N \}$ of cardinality $N(N=|F|)$.

- i. Apply GA to select the best subset of n_1 number of features from F to have the selected feature set as $F_1 = \{ f_1, f_2, f_3, \dots, f_{n_1} \}$ of cardinality $n_1 (n_1 = |F_1|)$ where $n_1 < N$.
- ii. Apply BFO to select the best subset of n_2 number of features from F to have the selected feature set as $F_2 = \{ f_1, f_2, f_3, \dots, f_{n_2} \}$ of cardinality $n_2 (n_2 = |F_2|)$ where $n_2 < N$.
- iii. Find the Intersection of n_1 features and n_2 features as n features as $F_3 = F_1 \cap F_2$.
- iv. Use the feature set F_3 for classification.

6. CLASSIFICATION

In pattern recognition and machine learning, classification is the problem of identifying which of a set of class a new observation belongs, on the basis of a training set of data containing observations (or instances) whose class is known. In this study the classification of the known faces from the unknown faces is implemented using SVM and BPNN. The SVM is a non-probabilistic binary linear classifier which classifies the test face input in to either known or unknown face. Another classifier used in this study is Back Propagation Neural network which is also a supervised classification technique like SVM.

Since the dimension of the feature space is high Gaussian kernel is used for classification with SVM. The Gaussian kernel is mathematically represented as

$$k(\mathbf{x}_i, \mathbf{x}_j) = \exp(-\gamma \|\mathbf{x}_i - \mathbf{x}_j\|^2),$$

for $\gamma > 0$.

Sometimes parametrized using $\gamma = 1/2\sigma^2$

In this study the eigenvectors are used as feature set and to reduce the dimension of the feature set the PCA and hybrid feature selection algorithm is used to select the best feature set. If the features set given for training the classifiers better describe the characteristics of the human face then the classification rate will be high.

7. EXPERIMENTS AND RESULTS

The experiments were performed using the sample frontal face images taken from the AT&T database. The database contains ten different images of each of 40 distinct subjects. All the images were taken in same lighting condition with the subjects in an upright frontal position. Here 5 different images of each of 20 distinct subjects were taken from the database. Ten images of a subject with homogeneous dark background extracted from the database are shown in Fig. 1.



Fig. 1 Images of an object in AT&T database.

Each image is of size 112 pixels by 92 pixels so therefore the size of the eigenvector will be 10304x1. Using the PCA technique 1000 eigenvector with highest Eigenvalue is extracted and then using the hybrid feature selection algorithm the dimension is further reduced. Feature sets containing varying length of eigenvector (50, 100, 200, 500) are selected. The classifiers are trained and tested with each feature set separately. On an average the ratio between the size of the initial feature set extracted from the sample images and the feature set selected using the hybrid feature selection algorithm is 100:2.

The classification accuracy is calculates using the formula

$$\text{Classification Accuracy} = (\text{Total No. of samples taken} - \text{No. of samples misclassified}) / \text{Total No. of samples taken}$$

The performance of the classifiers used in this study is described in the Table 1. In both the classifier the classification accuracy is higher when the feature set contains less number of eigenvector which has high eigen values. This shows that the eigenvectors with highest Eigenvalues will give better classification results when using the face images with homogeneous background and lighting conditions.

Table 1. Performance of the Classifiers

Size of Training set	SVMs' Classification accuracy	BPNNs' Classification accuracy
50	87.50%	88.00%
100	83.33%	84.00%
200	80.06%	81.10%
500	79.80%	80.20%

The Fig.2 & 3 compares the execution time required for the classification process for various size of feature vector. As the size of the feature vector increases the execution time is also increased linearly. The SVM requires less time for classification when compared to BPNN for all the sizes of feature vector.

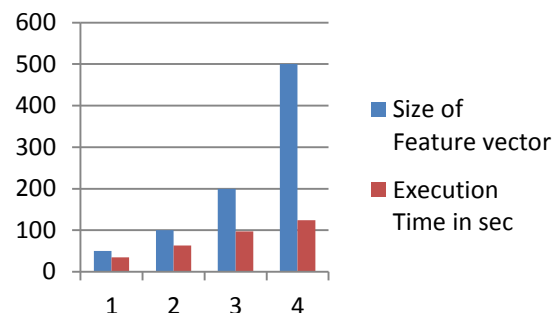


Fig 2. Comparison of Execution Time – SVM

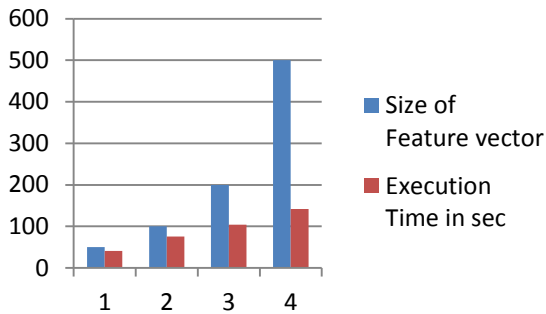


Fig 3. Comparison of Execution Time – BPNN

8. CONCLUSION

In this paper face recognition approach based on features selected using a hybrid meta-heuristic algorithm is proposed. The algorithm is applied to the feature vector extracted using Principal component analysis. Using the hybrid feature selection algorithm an optimal feature subset is selected from the feature space. The performance of the classifier is analyzed with various length of selected feature sets. The experimental results show that the even a very less number of eigenvectors with high Eigenvalues gives better classification rate when compared with the Eigenvectors with less Eigenvalues. Further this research will be extended to detect and recognize multiple faces or object in images using the hybrid-meta heuristic algorithms.

9. REFERENCES

- [1] Dimitri PISSARENKO. Eigen face –based facial Recognition. Dec 1, 2002.
- [2] Pramod kumar pandey, Yaduvir singh, Swetha tripathi. International Journal of Computer Applications (0975-8887). vol.15 No.4. February 2011.
- [3] Zakariya, S.M., Ali, R. Automatic Face reonition system by combining four individual algorithms. Computational Intelligence and Communication Networks (CICN) International Conference 2011, Pages: 222 - 226.
- [4] Cong Geng, Xudong Jiang. Face recognition using SIFT features. ICIIP'09 Proceedings of the 16th IEEE international conference on Image processing. Pages: 3277-3280 .
- [5] Donghyun kim, Hyeyoung park. An efficient face recognition through combining local features and statistical feature extraction. PRICAI'10 Proceedings of the 11th Pacific Rim International conference on trends in artificial intelligence. Pages: 456-466, Springer.
- [6] Hiremath, S., Joshi, D.G., Chadda, V.K., Bajpai, A. A multi-algorithmic face recognition algorithm. Advanced Computing and Communications, 2006. International Conference 2006, Page(s): 321 – 326.
- [7] C.-J. Tu, L.-Y. Chuang, J.-Y. Chang, and C.-H. Yang. Feature Selection using PSO-SVM. International Journal of Computer Science (IAENG), vol. 33, no. 1, IJCS_33_1_18.
- [8] E. Kokiopoulou and P. Frossard. Classification-Specific Feature Sampling for Face Recognition. Proc IEEE 8th Workshop on Multimedia Signal Processing, pp. 20-23, 2006.
- [9] Y. Yang, J. Wright, Y. Ma, and S. S. Sastry. Feature Selection in Face Recognition: A Sparse Representation Perspective, 2007.
- [10] X. Fan and B. Verma. Face recognition: a new feature selection and classification technique. Proc. 7th Asia-Pacific Conference on Complex Systems, December 2004.
- [11] D.-S. Kim, I.-J. Jeon, S.-Y. Lee, P.-K. Rhee, and D.-J. Chung. Embedded Face Recognition based on Fast Genetic Algorithm for Intelligent Digital Photography. IEEE Trans. Consumer Electronics, vol. 52, no. 3, pp. 726-734, August 2006.
- [12] M. L. Raymer, W. F. Punch, E. D. Goodman, L.A. Kuhn, and A. K Jain. Dimensionality Reduction Using Genetic Algorithms. IEEE Trans. Evolutionary Computation, vol. 4, no. 2, pp. 164-171, July 2000.
- [13] H. R. Kanan, K. Faez, and M. Hosseinzadeh. Face Recognition System Using Ant Colony Optimization-Based Selected Features. Proc. IEEE Symp. Computational Intelligence in Security and Defense Applications (CISDA 2007), pp 57-62, April 2007.
- [14] De Jong K. Learning with Genetic Algorithms: An overview. Machine Learning vol.3, Kluwer Academic publishers, 1988.
- [15] Haleh Vafaie and Kenneth De Jong. Genetic Algorithms as a Tool for Feature Selection in Machine Learning. Center for Artificial Intelligence, George Mason University.
- [16] Vafaie, H., and De Jong, K.A. Improving the performance of a Rule Induction System Using Genetic Algorithms. Proceedings of the First International Workshop on MULTISTRATEGY LEARNING, Harpers Ferry, W. Virginia, USA, 1991.
- [17] J. Adler (1966). *Chemotaxis in bacteria*. Science, vol. 153, pp. 708–716.
- [18] Passino K.M. Biomimicr of bacteria foraging for optimization and control. Control systems, IEEE. Vol.22 Issue 3. Page(s): 52-67.
- [19] H. Chen, Y.Zhu and R. Hu. Cooperative Bacterial Foraging Optimization. Research Article, Key Laboratory of Industrial Informatics, Shenyang Institute of Automation, Chinese Academy of Sciences, China, 2009.
- [20] Ahmed Al-Ani. Ant colony optimization for Feature subset selection. Proceedings of World Academy of Science, Engineering And Technology, February 2005, vol. 4 ISSN 1307-6884.
- [21] Rasleen Jakhar, Navdeep Kaur, Ramandeep Singh. Face Recognition Using Bacteria Foraging Optimization-Based Selected Features. (IJACSA) International Journal of Advanced Computer Science and Applications, Page(s): 106-111.
- [22] O. Ludwig and U. Nunes. Novel Maximum-Margin Training Algorithms for Supervised Neural Networks. IEEE Transactions on Neural Networks, vol.21, issue 6, pp. 972-984, Jun. 2010.
- [23] Z. Sun, X. Yuan, G. Bebis, S. Louis, Neural-network-based gender classification using genetic eigen-feature extraction, IEEE International Joint Conference on Neural Networks, May 2002.