ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 2, Issue 7, July 2014

FACE RECOGNITION USING EIGENFACE AND SUPPORT VECTOR MACHINE

MR. KANNAN SUBRAMANIAN[1]

MCA Department, Bharath Institute of Science and Technology, Bharath University, Chennai – 73[1]

ABSTRACT: Face recognition approaches faces many difficulties when there are variation in the face images due to lighting and other conditions. 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.

KEYWORDS: Meta-Heuristics, Genetic Algorithm, Bacteria Foraging Optimization, SVM.

I. 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. 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. In the proposed facial recognition system initially the sample face images are transformed in to Eigen faces using

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 2, Issue 7, July 2014

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]. Face recognition depend majorly on the human facial features such as the eyes, mouth, nose and face edges. A combination of four different inidivdual face recognition techniques was used by Zakariya S.M in [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 and 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. hybrid method for face recognition using local features and statistical feature extraction methods is proposed by Donghyun kim 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.

II. METHODOLOGY

Face Recognition (FR) is a matching process between a query face's features and target face's features. Face recognition (FR) has emerged as one of the most extensively studied research topics that spans multiple disciplines such as pattern recognition, signal processing and computer vision.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 2, Issue 7, July 2014

The modules can be further divided into image acquisition, feature extraction, feature selection and classifier.

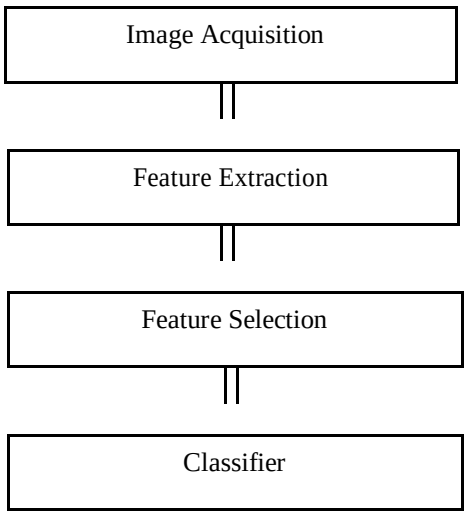

Figure 1. Overview of the proposed system.

A. Image Acquisition

Face images are acquired from the AT&T (Olivettti) Database (ORL) which are the variations in pose, scale, illumination, and facial expressions. Database of 50 face images are taken in same size. It has been done for accuracy in the result.

B. Feature Extraction

The first step in face recognition system is the extraction of the feature matrix. A typical feature extraction algorithm tends to build a computational model through some linear or nonlinear transform of the data so that the extracted feature is as representative as possible.

1) Eigenface Computation Using PCA

The Eigenface computation using PCA technique is briefly presented in this stage. 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. Let the training set images be $\Gamma 1, \Gamma 2, \ldots\ldots, \Gamma M$. 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. Finally find the Eigenvalues and Eigenvectors of the covariance matrix. If the size of the face image is M x N then the size of the eigenvector will be [M x N] x 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

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 2, Issue 7, July 2014

will also be higher. To avoid this eigenvectors with high Eigenvalues (i.e) the principal components are selected. characteristics of a individual face than the eigenvectors with less eigenvalue.

C. 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. Thus the Feature selection allows the reduction of feature space, which is crucial in reducing the training time and improving the prediction accuracy. In this work a hybrid approach using Genetic Algorithm combined with Information theory and Bacteria Foraging optimization is used to select the optimal features.

1) Feature selection using Genetic Algorithm

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. Each face image is transformed in to a vector of eigen-features. Therefore, the fitness evaluation contains two terms:

(i) accuracy and (ii) number of features used. The fitness function is calculated as proposed.

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

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.

2) Feature Selection using Bacteria Foraging Optimization

The BFO based feature selection method is used to which reduces the number of features, removes irrelevant, noisy and redundant data, and results in acceptable recognition accuracy. 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, Nc, Nre, n, where S : Number of bacteria in the colony, Nc: chemo tactic steps, Nre: reproductive steps, n: dimension of the search space

i. Initially place all the bacteria randomly at a position.

ii. For each reproduction step perform the following operations

iii. For each chemo taxis step perform the following operations means of corresponding classes and the grand mean in the feature space. Mi can be calculated as:

- Compute the fitness function using the following equation.
- Let w_1, w_2 ..., w_L and N_1, N_2... N_L denotes

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 2, Issue 7, July 2014

the classes and number of images within each class, respectively. Let M_1, M_2, M_L and M_0 be the

where $W^{(i)}_j$, j=1,2,…,Ni , represents the sample image from class W_i . The grand mean M_0 is computed using the formula

where N is the total number of images of all the classes. Thus the between class scatter fitness function F is computed as follows:

- Next change the bacteria to new position. 3.3 Compute the fitness function for the new position as in the first step. 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.

v. Calculate the health of each bacteria and Sort the bacteria in descending order of health.

vi.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.

vii. Pick up the position of bacteria B with max (health) value. This position represents the best feature subset of the features. 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.

3) Proposed Hybrid Algorithm

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

i. Apply GA to select the best subset of n1 number of features from F to have the selected feature set as

F_1={ f_1, f_2 ,f_3 … f_{n1}} 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 … f_{n2}} 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.

D. 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-

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 2, Issue 7, July 2014

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.

III. 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. Ten images of a subject with homogeneous dark background extracted from the database are shown in fig.1.

Each image is of size 112 pixels by 92 pixels. 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 .

Classification Accuracy = (Total No. of samples taken – No. of samples misclassified) / Total No. of samples taken.

The performance of the classifiers used in this study is described in the Table I. 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 I. PERFORMANCE OF THE CLASSIFIERS

Size of Training Set	SVMs' Classification Accuracy	BPNN's Classification accuracy
50	87.50%	88.00%
100	83.33%	84.00%
200	80.06%	81.10%
500	79.80%	80.20%

Figure. 3 and 4 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.

IV. CONCLUSION

The face recognition approach based on features selected using a hybrid meta-heurisitc algorithm is

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 2, Issue 7, July 2014

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 lengths 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 system will be extended to detect and recognize faces or object in images using local face representations, Scale Invariant feature transform(SIFT), Local Binary Pattern(LBP).

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 reconition system by combining four individual algorithms. Computational Intelligence and Communication Networks (CICN) International Conference 2011, Pages: 222 - 226.

[4] 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.

[5] 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.