A

dissertation

on

TAXONOMY OF NATURE INSPIRED ALGORITHM FOR FACE RECOGNITION

Submitted in partial fulfilment of the requirement

For the award of the degree of

MASTER OF TECHNOLOGY

IN

COMPUTER SCIENCE & ENGINEERING

By

Mani Gupta

(Roll No. 2K13/CSE/10)

Under the guidance of

Prof (Dr.) Daya Gupta

Department of Computer Engineering

Delhi Technological University, Delhi

Delhi Technological University, Delhi

2013-2015

DELHI TECHNOLOGICAL UNIVERSITY

CERTIFICATE

This is to certify that the project report entitled **"Taxonomy of nature inspired algorithm for face recognition"** is a bonafide record of work carried out by **Mani Gupta (2K13/CSE/10)** under my guidance and supervision, during the academic session 2013-2015 in partial fulfilment of the requirement for the degree of Master of Technology in Computer Science & Engineering from Delhi Technological University, Delhi.

Dr. Daya Gupta

(Project Guide)

Department of Computer Engineering

Delhi Technological University

Delhi-110042

DELHI TECHNOLOGICAL UNIVERSITY

ACKNOWLEDGEMENT

With due regards, I hereby take this opportunity to acknowledge a lot of people who have supported me with their words and deeds in completion of my research work as part of this course of Master of Technology in Computer Science & Engineering.

First of all, I would like to express my deepest sense of gratitude and indebtedness to my guide and motivator, **Dr. Daya Gupta**, Professor, Department of Computer Engineering, Delhi Technological University for her valuable guidance and support in all the phases from conceptualization to final completion of the project.

I humbly extend my grateful appreciation to Ms. Kanishka Bansal who has been an excellent motivator throughout the work. Lastly,I am grateful to all my friends and family for their continued support and encouragement throughout the research work.

Mani Gupta

Roll No. 2K13/CSE/10

M.Tech (Computer Science & Engineering)

Department of Computer Engineering

Delhi Technological University

Delhi - 110042

ABSTRACT

Face Recognition is one of the crucial areas of research because of its widespread application. It is also a real world problem that can be solved with the help of computational intelligence techniques . Now, days area of computational intelligence is gaining a lot of interest. Many of the nature –inspired methodologies and approaches comes under this area are used to solve the real-world problem to which traditional approaches are infeasible, ineffective and less efficient. Computational intelligence primarily includes artificial neural networks, evolutionary computation, swarm intelligence and fuzzy logic.

Face Recognition is a two step process i.e. features extraction and recognition process. In the feature extraction phase Gabor kernel is used to smoothen the images and PCA is used for feature extraction. Later on an evolutionary algorithm is used to find optimal features and an evolutionary algorithm is used to recognize an input image. We worked on both the phases to increase efficiency, by applying various evolutionary techniques.

There are so many techniques available to solve the problem that the application developer gets into dilemma that which technique is most appropriate to use. In our thesis we have tried to solve the problem of face recognition using some of the evolutionary techniques like Ant Colony Optimization(ACO), Particle Swarm Optimization(PSO) ,Hybrid ACO/PSO, Biogeographical Based Optimization(BBO), Extended BBO and holistic technique like Principal Component Analysis(PCA). We analyzed and compared the results of all of these techniques to make it clear that which one is appropriate.

Performance analysis is performed using Olivetti research Laboratory (ORL) face database and Cohn-Kanade database. We have shown the performance on the basis of time taken for recognition and accuracy in recognition. We found that different technique are best on both the parameters but if we are trying to find the appropriate one, Extended BBO serves the purpose.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Chapter 1

Introduction

Since 1960’s ,there is a lot of research work going on in the field of face recognition still face recognition is an active area of research. Face of each human has some unique characteristics and face recognition is technique that uses those characteristics for identification. Face recognition has become a popular area of research in computer vision. Due to the availability of feasible technologies for face recognition, it has become one of the most successful applications of image analysis and understanding. The problem statement of face recognition application can be stated as to match a given input image against a large database of images and determine whether it is there or not. Face recognition is usually used for two purpose identification (one to one matching) and verification (one to many matching). Based on these operations it is used in various applications like national security, Surveillance, drivers licence verification, General identity verification tools, “Smart Card” applications, human computer interface and many other [22].

Now, days area of computational intelligence is gaining a lot of interest. Many of the nature – inspired methodologies and approaches comes under this area are used to solve the real-world problem to which traditional approaches are infeasible, ineffective and less efficient. Computational intelligence primarily includes artificial neural networks, evolutionary computation, swarm intelligence and fuzzy logic. It is of great usage in the areas of image processing, data mining etc. Face recognition is also a real world problem that can be solved with the help of computational intelligence techniques. There are so many techniques available to solve the problem that the application developer gets into dilemma that which technique is most appropriate to use. In our thesis we have tried to solve the problem of face recognition using some of the evolutionary techniques like Particle Swarm Optimization(PSO) ,Ant Colony Optimization(ACO), Hybrid ACO/PSO, Biogeographical Based Optimization(BBO), Extended BBO and holistic technique like Principal Component Analysis(PCA). We analyzed and compared the results of all of these techniques to make it clear that which one is appropriate.

1.1 Motivation

Face recognition methods can be categorize into feature based and holistic method and hybrid of these two methods for intensity images. One of the methods of holistic technique is Principal Component Analysis. Many techniques are present to extract features from an given image like DCT,DWT, PCA but PCA ia a standard and most common technique.. It is use to extract desirable feature from an image. Principal Component Analysis (PCA) is an eigenface algorithm that uses dimensionality reduction to find the vectors which best account for the distribution of face images within the entire image space.

Ant Colony Optimization(ACO) is an evolution simulation algorithm proposed by M. dorigo et al[11].It is inspired by ant's social behavior. It was first used for solving Travelling salesman problem and now it has been successfully used for system fault detecting, job-shop scheduling, network balancing, robotics, graph coloring and many other combinational optimization problems. ACO is used for feature selection after the feature extraction phase[16]. This involves heuristic and random search stratergies in order to reduce the computation time for face recognition. Particle Swarm Optimization (PSO) is an evolutionary computational paradigm created by Kennedy and Eberhart [28]. It is a population based optimization algorithm where each particle is an individual and the swarm is composed by particles. The problem solution is formulated as a search space and each position at that space is a correlated solution of the problem. PSO is used in face recognition to reduce the execution time and increase the accuracy to identify an image. This motivated to develop a hybrid approach using ACO and PSO to solve the problem of face recognition.

Biogeography is the study of the geographical distribution of biological organisms. Engineers have a mindset that we can learn from nature. Like mathematics of genes inspired the development of genetic algorithm, mathematics of biogeography lead to the development of a new algorithm: Biogeographical based optimization. It was first presented by D. Simon in year 2008 [29].As BBO has some features common with GA and PSO, it can be applied on the real world problems on which GA and PSO can be applied i.e high dimension problems with multiple local optima. This motivates the application of biogeography to optimization problems. Just as the mathematics of biological genetics inspired the development of genetic algorithms (GAs), and the mathematics of biological neurons inspired the development of artificial neural networks, this paper considers the mathematics of biogeography as the basis for the development of a new field: biogeography-based optimization (BBO). We discuss natural biogeography and its mathematics, and then discuss how it can be used to solve optimization problems. However, BBO also has some features that are unique among biology-based optimization methods. Simon[29] demonstrate the performance of BBO on a set of 14 standard benchmarks and compare it with seven other biology-based optimization algorithms. We also demonstrate BBO on a real-world on real world sensor selection problem for aircraft engine health estimation. This all motivated us to apply BBO on the problem of face recognition.

Extended Species Model of Biogeography is an extension of BBO. It involves an additional dependency factor that shows interdependence on various species over each other as in predator prey relationships. It also involves factor of relevance that states if a particular SIV is greater than some threshold, than only it should be considered in process of evolving. It also involves decline and growth rate that are function of extinction rate, emigration rate, evolution rate and immigration rate [13]. Therefore it led a modification in Simon model. This is the essence of extended BBO.

1.2 Related Work

Face recognition has applications in various areas but it is mainly important in security system. In today's world, automatic face recognition is one of the utmost requirements. In face recognition process, our first step is to extract features; this is one of the fundamental steps. Researchers are looking for the most efficient face recognition techniques. There are various approaches to solve this problem like Linear Discriminant Analysis [4], Discrete Cosine Transform [1], Principal Component Analysis [9, 24, 32]. Semi supervised methods [4] and hybrid approaches involving neural network, are proposed by many researchers. They gave a combination of supervised learning as well as unsupervised learning, at features extraction phase unsupervised learning and at recognizing phase supervised learning, so classification error will be reduced. These algorithms can't be applied alone on raw images to act efficiently, they need some other perquisite conditions for example Principal Component Analysis requires proper alignment of images as well as constant illumination. Bhuian, er al. 2007 opted for the combination of these algorithms to form hybrid approaches for face recognition. Hybridization of Gabor filters with neural network was done to achieve higher accuracy [4]. Chung, et al., 1999 have clubbed Principal Component Analysis with Gabor Kernel to achieve high accuracy [6]

Recently some researchers have used natural inspired techniques like Ant Colony Optimization; Particle Swarm Optimization and hybrid PSO/ACO. They have shown amazing results [16,33, 20]. Recently back a new nature inspired technique, Biogeography based Optimization is been given by Dan Simon [29]. BBO has been applied to face recognition on both principal component analysis [8] and Discrete cosine transform. Very recently Extended Species Model was proposed by Lavika Goel [13], this was an extension of BBO, has been used in remote sensing for edge detection, this model outperform BBO. This motivated us to adapt all these algorithms for face recognition. We tried to apply all these algorithms in the area of face recognition and compare and analysed them.

1.3 Problem Statement

The purpose of face recognition system is to match a given input image against a set of database consisting of a number of images and to find out whether the image matches or not. There are a number of evolutionary techniques available for this purpose. Since face recognition is related to security system, we need to find an appropriate algorithm out of many to serve this purpose.

Face recognition has two phases: feature extraction and image recognition. Feature extraction phase is also followed by feature selection. Since there are so many techniques available to solve the problem of face recognition that the application developer gets into dilemma that which technique is most appropriate to use. In our thesis we have tried to solve the problem of face recognition using some of the evolutionary techniques like Particle Swarm Optimization(PSO) , Ant Colony Optimization(ACO), Hybrid ACO/PSO, Biogeographical Based Optimization(BBO), Extended BBO in feature selection and recognition phases and used holistic technique like Principal Component Analysis(PCA) in feature extraction phase . We analyzed and compared the results of all of these techniques to make it clear that which one is appropriate.

So thesis problem can be stated as follows:

"**To study the taxonomy of various nature inspired algorithms for face recognition**".

1.4 Scope Of The Work

In this work we have compared various nature-inspired algorithms for Face Recognition. In a face recognition model we have training phase and the recognition phase. The training phase is divided into three steps. In the first step we filter the images with gabor kernel to align and smoothen them .In the second step we applied PCA to extract features (i.e dimension reduction) and in the last step we used different swarm intelligence algorithms(ACO,BBO, Extended BBO) to find optimal features amongst the output of PCA. At the recognition phase also we have used different swarm intelligence algorithms (PSO, BBO, Extended BBO) to recognize the input in less amount of time.

Different combinations of swarm intelligence algorithms that we have tried are as follows:

A. Particle Swarm Optimization (PSO) in recognition phase

In training phase we have two steps. Firstly the gabor kernel is used then PCA is used to extract the features. In the recognition phase we used particle swarm optimization to reduce the time consumed for identification of a test image.

B. Ant Colony Optimization (ACO) in training phase

In training phase we have three steps. Firstly the gabor kernel is used then PCA is used to extract the features and in the last step we use ant colony optimization for feature selection to get a optimal feature vector set. In the recognition phase we simply use minimum distance classifier to recognize the input test image.

C. ACO in training phase and PSO in recognition phase (Hybrid ACO/PSO)

In the training phase we have three steps: first we filter images with gabor kernel to align images and smoothen them. In the second step we applied PCA to extract features and in the last step we use ant colony optimization for feature selection to get a optimial feature vector set. In the recognition phase we used particle swarm optimization to reduce the time consumed for identification of an input test image.

D. Biogeographical Based optimization (BBO) in training and recognition phase

In the training phase we have three steps In the first step we filter images with gabor kernel to align images and smoothen them. In the second step we applied PCA to extract features and in the last step we use biogeography based optimization for feature selection to get a optimial feature vector set.In the recognition phase we have used biogeography based optimization to recognize an input test image.

E. Extended BBO in training and recognition phase

In the training phase we have three steps. In the first step we filter images with gabor kernel to align images and smoothen them. In the second step we applied PCA to extract features and in the last step we use BBO based on extended species model of Biogeography for feature selection to get a optimal feature vector set. At recognizing phase we have used BBO based on extended species model of Biogeography to recognize an image. In this we input an image and apply algorithm to find a best match without comparing our input image with all the images in training database. This led to reduction in time and makes it an efficient approach.

We tested our approaches on ORL face database acquired at the Olivetti Research Laboratory, Cambridge, U.K and Cohn-kanade AU- coded facial expression by Kanade, T.Cohn, J.F., & Tian, Y.(2000) The results obtained prove the correctness of our approach.

The scope of the work is described as:

TRAINING PHASE		RECOGNITION PHASE
FEATURE EXTRACTION	FEATURE SELECTION	
Gabor filter + PCA	None	PSO
Gabor filter +PCA	ACO	Minimum distance classifier
Gabor filter +PCA	ACO	PSO
Gabor filter +PCA	BBO	BBO
Gabor filter + PCA	Extended BBO	Extended BBO

Table 1 Various face recognition algorithms

We also compared the efficiency of various algorithms with our approach. The training accuracy is shown on the basis osf ROC curves . We compared these algorithms on the

- basis of time taken for recognition
- basis of accuracy in recognition.

1.5 Organization of the Dissertation

Rest of work is organized as follows:

Chapter 2: **Face Recognition** : In this chapter we explain the process of face recognition . It is a three step process consisting of face detection, feature extraction and face recognition or classification.We will give in this chapter an overview of principal component analysis and gabor kernel and minimum distance classifier.

Chapter 3: Generic Evolutionary Techniques: In this chapter we will explain various nature inspired techniques like Paricle swarm optimization(PSO), Ant Colony Optimization(ACO), Hybrid of ACO and PSO, Biogeographical Based optimization, Extended Biogeographical Based Optimization.

Chapter 4: Computational Intelligence in Face Recognition: In this Chapter we will adapt various evolutionary algorithms like ACO,PSO, Hybrid ACO/PSO, BBO, Extended BBO in face recognition. We will also explain how these techniques are used in the process of face recognition and gives the algorithm for each for recognizing a face

Chapter 5: Experimentation, Evaluation and Results : In this chapter we will compare all the evolutionary techniques explained in previous chapter on the basis of recognition time and accuracy. We have tested and shown our results on two datasets.

Chapter 6: **Conclusion**: This chapter includes the conclusion of the thesis and the future scope of work.

References

Appendix- A Abbreviations

Appendix-B Matlab

Chapter 2

Face Recognition

In this chapter we explain the process of face recognition . It is a three step process consisting of face detection, feature extraction and face recognition or classification. We will also give in this chapter an overview of principal component analysis and gabor kernel and minimum distance classifier.

2.1 Steps of face recognition

1. Face Detection
2. Feature extraction and selection
3. Face recognition

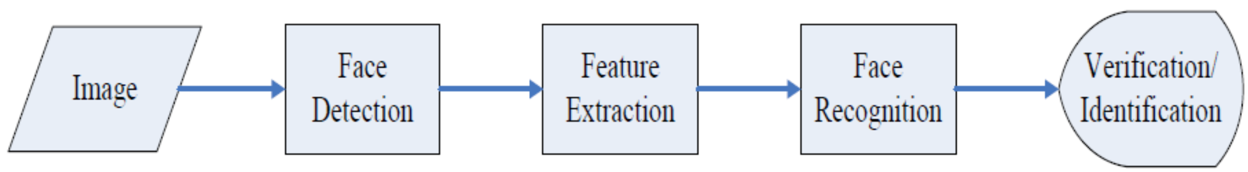

Figure 1: steps of process of face recognition

2.1.1. Face Detection:

The main function of this step is to determine whether human faces appear in a given image and if yes, where these faces are located at[2]. The expected outputs of this step are patches containing each face in the input image. In order to make face recognition system more secure and easy to design,we perform **face alignment** to justify the scales and orientations of these patches. Apart from serving as the pre-processing for face recognition, face detection could be used as an area of interest of detection, retargeting, video and image classification, etc.

2.1.2 Feature Extraction and Selection:

After the face detection step, human-face patches are extracted from images. We cannot use these patches directly for feature .First, each patch usually contains over 1000 pixels, which are too large to build a robust recognition system. Second, face patches may be taken from different camera alignments, with different face expressions, illuminations, and may suffer from occlusion and clutter. To overcome from all these drawbacks, feature extractions are performed to do information packing, dimension reduction, salience extraction, and noise cleaning.

After this step, a face patch is usually transformed into a **vector with fixed dimension** or a set of **fiducial points and their corresponding locations**. In some literatures, feature extraction is either included in face detection or face recognition[2]. The basic responsibility of this module is to compose a feature vector that is enough to represent an image. In our proposed work we have used Principle Component Analysis for the feature extraction phase

Feature selection seeks for the optimal set of d features out of the m features obtained from feature extraction module. Many algorithms have been proposed for the feature selection. In our work we have used Ant colony optimization for performing feature selection.

2.1.3 Face Recognition:

After extracting the features from each image and making a representation from it, the last step is to recognize the identitiy of a test face. In order to achieve automatic recognition, a face database is required to build. For each person, several face images are taken and their features are extracted and stored in the database. Then when an input face image comes in, we perform face detection and feature extraction, and compare its feature to each face class stored in the database[2].

There are two general applications of face recognition, one is called **identification** and another one is called **verification**. Face identification means given a face image, and now we expect from the face recognition system to tell the probable identification that with whom this face is similar; while in face verification, given a face image and a guess of the identification, we want the system to tell true or false about the guess. In fig. 2, we show an example of how these three steps work on an input image.

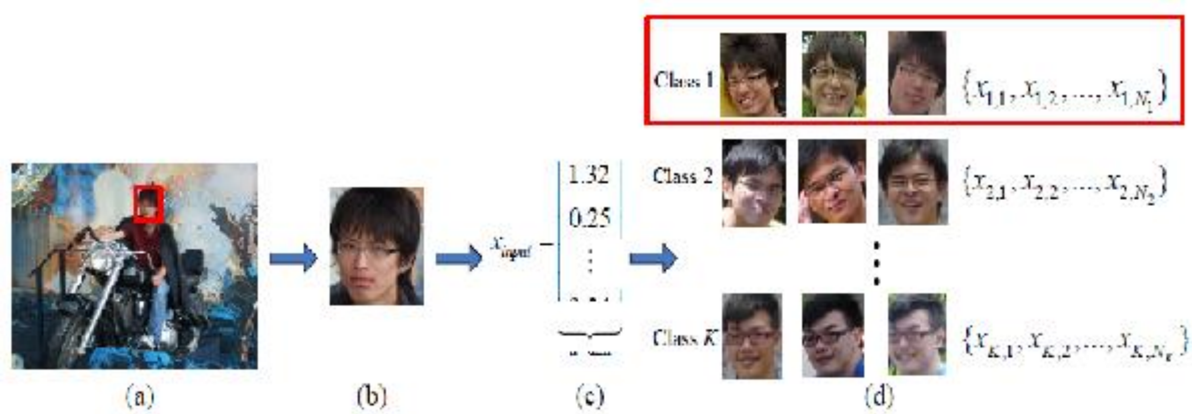

Figure 2: An example of how the three steps work on an input image. (a) The input image and the re-sult of face detection (the red rectangle) (b) The extracted face patch (c) The feature vector after feature extraction (d) Comparing the input vector with the stored vectors in the database by classification techniques and determine the most probable class (the red rectangle). Here we express each face patch as a *d*-dimensional vector, the vector as the, and as the number of faces stored in the k_{th} class.

From the above these phase mentioned we are concerned only about the last two phases i.e the feature extraction and selection ,and face recognition. Now we will explain each phase in detail with the techniques that can be used in these phases.

.

2.2 Principal Component Analysis (PCA) :

Face recognition algorithms can be divided into two broad categories according to feature extraction schemes of face recognition:

Principal component analysis is basically used for dimensionality reduction. The main task of PCA is to analyse the data and to find and identify the patterns in it. Finding out the patterns in dataset helps us to reduce the dimentionality of dataset with minimal loss of information.

Principal component analysis is basically used to project a feature space of higher dimensionality onto a smaller subspace that represents our data "well". The motive behind reducing the dimensionality is to reduce the computational cost needed to evaluate the features[32].

Principal Component Analysis is a statistical procedure that illustrates the covariance structure of a set of variables. In particular it allows us to identify the principal directions in which the data varies. For example, in figure 1, we represented two variable dataset in X-Y coordinate system. Using PCA we found out the principal directions in which the data varies i.e. the U-axis. The second important we have V-axis orthogal to U-axis. If we place the $U - V$ axis system at the mean of the data it gives us a compact representation . Now we map each point of X-Y coordinate system into corresponding U-V system. Mapping the data points ,makes the data de-correlated,means the co-variance between U and V becomes zero.

For a given dataset, PCA finds the principal components. These peinciple components are the axis or the principal directions along which the data varies. In figure3 U-V axis system are the principal components.

As we see in Figure 3, there are two principal axis U and V. We can also see that V component of all the data points are close to zero. This V component might have come in data points due to some external factor for example due to noise. Thus in the U-V axis system we can represent the data set by one variable U and discard v. This is how we reduce the dimensionality of the data set. And hence the dimensionality in this case is reduced by 1.

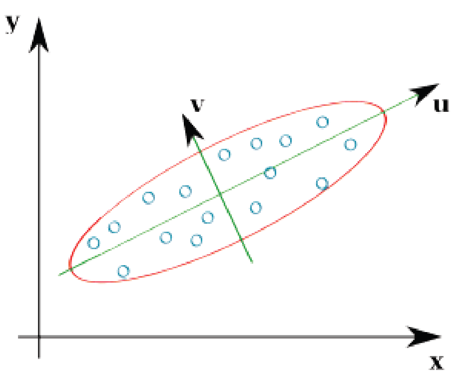

Figure 3 : PCA for data representation

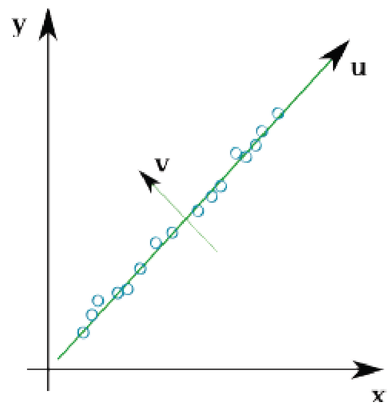

Figure 4: PCA for dimension reduction

2.2.1 PCA Algorithm

1) Take the entire dataset consisting n samples each of having the dimension d making d*n matrix
2) Compute the d-dimensional mean vector of the d*n matrix
3) Compute the co-variance matrix of the d*n matrix (i.e of the entire dataset)
4) Calculate the eigenvectors and corresponding eigen values of the covariance matrix obtained.
5) Sort the eigen vectors by decreasing eigen values and choose k eigen vectors of higher eigen values. Each eigen vector is of d-dimension, hence forming a d*k dimensional matrix W.
6) Use this d x k eigenvector matrix to transform the samples onto the new subspace. :

   $$y=W^T*x \tag{1}$$

   (where **x** is a d x 1 -dimensional vector represents one sample, and y is the transformed **k** x 1 -dimensional sample in the new subspace.)

2.2.2 PCA in face recognition

Feature extraction is one of the important steps in face recognition. PCA is one of the techniques used in feature extraction. The image in considered to be in a form of N x N matrix, which can be expressed as a set of points in the space N x N dimensions. The aim of PCA is to find out the required vectors that can be used to represent the image information.

PCA algorithms steps[8] are as follows:

1) Take m images in input of N * N and convert them into gray scale. Convert each image into a $N^2 *1$.
2) Calculate the mean face of $N^2 *1$. Substract mean face from each of the face image and store it in a matrix B.
3) Calculate the covariance matrix using $C=BB^T$ of N^2*N^2. But since this is a higher dimension matrix we will find the covariance matrix for computation using B^TB.
4) Calculate the eigen values and corresponding eigen vectors of covariance matrix. For m images we will get corresponding m eigen vectors like $v_1,v_2\ldots\ldots\ldots v_m$. The eigen vectors of actual covariance matrix is found by $u_i=Bv_i$ $(1<i<=m)$.Store these eigen vectors $u_1,u_2,u_3\ldots\ldots\ldots\ldots\ldots u_m$.
5) Find out k eigen vectors out of m eigen vectors corresponding to higher eigen values.
6) All the images calculate their weight matrix corresponding to eigen vectors representing their proportion in them by iterating $W_j = u_j^T *Q_i$ where $j=1,2\ldots\ldots k$ and $i=1,2\ldots\ldots m$. For each image we have a weight matrix of $K*1$.

2.3 Minimum Distance Classifier

Minimum distance classifier is the simplest classifier which is based on minimum error calculation. Error is calculated as norm or Euclidean distance

Say ‘A’ and ‘B’ are 2 vectors of dimension k.

$$\text{Norm}(A,B) = \|A-B\| = \text{Euclidean distance} = \sqrt{\sum_{i=1}^{k}(A^i - B^i)^2} \qquad (2)$$

Suppose we are given m training examples as (x^1, y^1), (x^2, y^2), ... (x^m, y^m) .x is input and y is the output class. And there is an unknown test sample X_i. We have to determine to which class this test sample belongs to. For this we will calculate distance of X_i from each training sample .distance if calculated as norm of the vector. And then we will select that training sample which is closest or has minimum error or has minimum Euclidean distance. The output label y of the selected training sample will be our class.

2.4 Gabor Filter

Gabor filter also known as Gabor wavelets or Gabor kernel. They are also known as orientation sensitive filters. They travel in packs, anf if set in one direction they give strong reponse for the target images structure in that direction.They are complex band pass filters. Their shape resembles with the visual cortex of the mammalian brain. There applications include spatially local features extraction of a confined frequency band, extraction of multi-resolution, feature tracking [7].

Another important use of gabor filter is it is a very efficient tool for facial feature extraction.In spatial domain the set of 2D Gabor kernel can be expressed as [6]

$$\Psi_{u,v}(a', b') =(f^2u/ \pi hp)e^{-l}e^{j2\pi tfu} \qquad (3)$$

Where

$t = a'\cos\Phi v + b'\sin\Phi v$

$d = -a'\sin\Phi v + b'\cos\Phi v$

f_u = fmax/2(u/2) and $\Phi v = v\pi/8$

$l=((f^2u/k^2)t^2+(f^2u/p^2)d^2)$

The expression of gabor kernel represents a gaussian filter function by a complex plane wave. Central frequency and orientation can be represented by fu and θv, respectively. And h and p evaluates ratio between central frequency and the size of Gaussian envelope.Though we can also have different values for the above mentioned parameters evaluating characteristics of the kernels, the most common parameters which is used for the face recognition are $\kappa = \eta = \sqrt{2}$ and fmax = 0.25 [32,24].

Chapter 3

Generic Evolutionary Techniques

In this chapter we will explain various nature inspired techniques like Paricle swarm optimization (PSO), Ant Colony Optimization (ACO), Hybrid of ACO and PSO, Biogeographical Based optimization, Extended Biogeographical Based Optimization.

3.1 Particle Swarm Optimization

Particle swarm optimization (PSO) is a population based stochastic optimization technique inspired by social behavior of bird flocking. It was developed by Dr.Eberhart and Dr.Kennedy in 1995[28]. In PSO, system is initialized by a population of particles representing random solutions. From these random solutions we find a candidate solution and our task is to optimize that candidate solution iteratively generation by generation. The particles moves in the search space and the position and velocity of particles are updated in each generation. In simple words PSO simulates the behavior of bird flocking.

For example, a group of birds searching for food in a search space and there is only one piece of food in search space. No bird knows where the food is, but they know how far the food is in each iteration. The birds adopts the strategy to follow the bird which is nearest to the food. The position and velocity of each bird is updated in each iteration.

In PSO, each single solution instead of “bird” is called “particle”. Fitness value of each and every particle is evaluated in each iteration. The particle having the best fitness value said to be global best or gbest and its fitness value as gbest value. The velocity of the particle that is updated in each iteration depends upon the gbest value and the personal best value(pbest) of the particle itself. . When a particle takes part of the population as its topological neighbors, the best value is the lbest.

After finding the two best values gbest and pbest , the particle updates its velocity and positions with following equations(4) and(5) :

1. ***v[]=v[]+c1*rand()*(personalbest[]-current[])+c2*rand()*(gbest-current[])*** (4)

2. ***current[]=current[]+v[]*** (5)

- v[] is the particle velocity
- persent[] is the current particle (solution).
- pbest[] and gbest[] are personal best value of the particle and gbest is the global best value among all the particles.
- rand () is a random number between (0,1). c1, c2 are learning factors, usually c1 = c2 = 2.

3.1.1 General Algorithm for PSO:

1) Intialise all the particles with random position and velocities.
2) Calculate the fitness of each particle using a fitness function. If the fitness of the particle is better than its previous fitness value then update the personal best fitness value.
3) After calculating the fitness values of all the particles, update the global best fitness value with the best among all the particles. And the position of the particle having global best fitness value, is taken as global best position.
4) Update the positions and velocities of all the particles using personal best position and global best position.
5) We repeat the process again from step 2 if maximum iterations is not reached.

3.2 Ant Colony Optimization

Ant colony optimization algorithm **(ACO)** is a probabilistic technique for solving optitimization problems that can be mapped in finding good paths in a graph[11].

ACO was firstly used for solving travelling salesman problem. To solve a new problem by ACO, we try whether that problem can be think of or mapped in travelling salesman problem.

Motivation- ACO algorithm is motivated from intelligent behaviour of ants in finding the almost optimal path from their source to destination. Ant lay down pheromone trail while moving in serch of food .other ants follow that pheromone trail .On the basis of amount of pheromone in a trail ants choose the path. They might start their journey on random paths ,but eventually all of them reach to their destination choosing the optimum or nearly optimum path.

ACO was initially proposed by Marco Dorigo in 1992 in his PhD thesis. The general aco solves the TSP as follows-

3.2.1 Tsp Using Ant Colony Optimization

TSP is a combinatorial problem, which takes large amount of time to solve by naïve approach of checking all possible tour combination. Ant colony optimization can find a tour which is approximately close to optimal tour.

The number of ants are fixed to some constant say 10 and each ant is allowed to construct a tour picking up a random city at the beginning. Then from that city, i the ant picks one of the unvisited neighbour cities, j with the probability. After all the ants complete the tour the pheromone update is done and also during this update, the evaporation of pheromone is taken into consideration. Once this is done, the ants are allowed to construct the tours again and this continues.

Parameters And Terms Used In the algorithm:

P denotes the evaporation factor of pheromones which is is a constant.

1) **ALPHA** (α) denotes the weightage given to pheromones while deciding the path to be taken probabilistically
2) **BETA** (β) denotes the weightage given to heuristic value while deciding the path to be taken probabilistically.
3) **Heuristic Value** (ϑ) is the inverse of the cost of respective edge.
4) **Pheromone Initialization:** The pheromone value of all edges are initialized to $\frac{1}{\#\, of\ cities}$
5) **Pheromone Update:**

$$T_{ij} = T_{ij} + \Delta\, T^{k}_{ij} \qquad (6)$$

$$\Delta\, T^{k}_{ij} = \begin{cases} \frac{Q}{l_k}, & if\ k^{th}\ ant\ picked\ i,j\ edge \\ 0, & otherwise \end{cases}$$

Where l_k is the length of tour by ant k and Q is a constant.

6) **Pheromone Evaporation:**

$$T_{ij} = T_{ij} * (1- P) \tag{7}$$

7) **Probability Of Traversing An Edge:** $\frac{(T_{ij})^{\alpha}+(\vartheta_{ij})^{\beta}}{\Sigma_k (T_{ij})^{\alpha}+(\vartheta_{ij})^{\beta}}$ (8)

3.3 Biogeographical Based Optimization:

Biogeography based optimization is a nature inspired algorithm based on the geographical distribution of biological organisms. In 1960s, the mathematical model of BBO was given by Robert Mac Arthur and Edward Wilson. The model explained how the species distributed in the neighbouring islands. It basically demonstrated how and when the species migrate from one island to the other, how species become extinct and how species arises.

An island is a place where the species resides so it is called as 'habitat'. So instead of island we will use the term habitat. Geographical areas that are well suited for residence of the species have high habitat suitability index(HSI). Factors that lead to high HSI of a particular habitat are rainfall, diversity in vegetation, temperature etc[29].

The variables that define habitability are known as Suitability index variables (SIVs). SIV s are the independent variables but the HSI of each habitat depends on SIVs.Therefore habitats with high HSI value has large number of species and habitats with low HSI value has less number of species.

Now we explain the relation between the emigration rate and immigration rate in high and low HSI habitats. Emigration is when a species move leaves its own habitat and immigration is when a species comes in habitat. High HSI habitats tends to have high emigration rate of species towards other habitats due to the large number of species in it. On the other hand high HSI habitats have low immigrate rate of species as they have already achieved the saturation in number of species. Habitats with low HSI have high immigration rate as they have low population in them. But as the species move in towards the habitats with low HSI the HSI of those habitats increase because greater is the population in the habitat it corresponds it is more suitable for residence. If HSI of a habitat remains low, then the species tends to go extinct, which will further add on to additional immigration. So, the habitats with high HSI are more static in comparision to the habitats with low HSI.

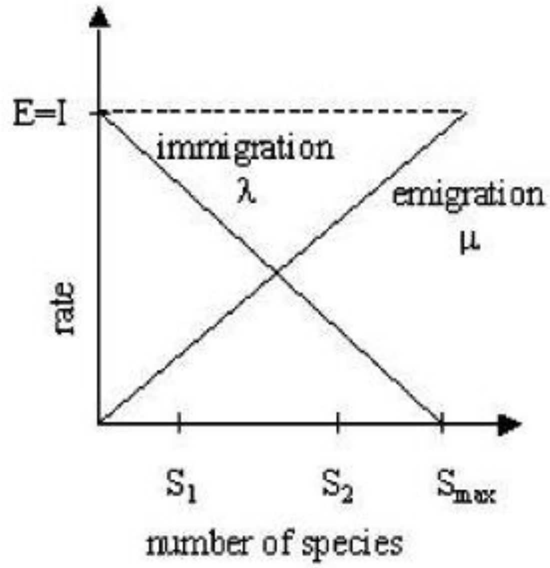

Figure 5 represents species abundance in Single Island.

We plot a graph as shown in Figure5 between the rates i.e the immigration and emigration rate and the number of species in the habitat. The graph plots the curves of immigration and emigration rates on the basis of numberof species in the habitat. As we can see in the figure $S_1 < S_2 < S_{max}$ corresponds to the number of species in habitat.

When the number of species is zero in habitat. The immigration rate is maximum and the emigration rate is zero. As the number of species increases, the curve of immigration rate declines and the curve of emigration rate rises.When number of species are between S_1 and S_2 there exist an equilibrium between both the rates. As the number of species reaches to its maximum value S_{max} the emigration rate is at its maximum and the immigration rate becomes zero as the number of species has reached its saturation.

Migration and Mutation are two main process in this algorithm[8]. These two are explained below:

MIGRATION:

When we have a candidate solution, we evaluate the fitness of the solution using a fitness function. Solution with high HSI is considered good and solution with low HSI is bad.

A candidate solution can be expressed as array of integers and each integer corresponds to a SIV.

Migration has two processes emigration and immigration. . Emigration and Immigration rates of each solution share information probabilistically between habitats. We modify habitat based on other solution. If a particular solution is selected to modify then we will find out which SIV of that solution is to be moved or immigrated. We will calculate the immigration rates of all the SIVs in that solution and sort them in decreasing order. Siv having the highest immigration rate is to be immigrated. Now we will also have to find out that which SIV is to be emigrated in place of immigrated SIV from other habitats.

Procedure of migration:

1) Select a solution for modification based on the probability proportional to the immigration rates(λ_i).
2) One of the SIV from the habitat chosed in immigrated into some other habitat, let this habitat be k. But before that a habitat is chosen whose SIV will be emigrated in this habitat on the basis of emigration rates(μ_i) and let this chosed habitat be m.
3) Now we randomly select one SIV from m and replace a SIV of k with it.

MUTATION:

A habitat's HSI value can be randomly changed due to certain events. This random change in HSI value can be expressed in terms of mutation rates. Mutation depends upon the species count probabilities of the habitat. Low species count in a habitat and high species count in a habitat both have low probabilities and medium species count have high probability. A solution having high probability has less chances of mutation to other solution.

1) We calculate the probability(P_i) of a habitat based on immigration rate(λ_i)an emigration rate(μ_i).
2) Select a SIV from the habitat proportional to the probability $P_{i.}$
3) After selecting a particular SIV replace this with a randomly generated SIV.
4) Iterate it for all the SIV's.

3.4 Extended Species Abundance Model Of Biogeography (Extended BBO):

Extended BBO was proposed by [6], it was an extension of BBO. BBO was based on emigration and immigration rates. Emigration is leaving one's own native place and immigration is arrival of a species to new place .

In extended BBO a new factor is included i.e. dependency factor with emigration rate. It is important in finding emigration rate as species tends to resist migration due to dependency on one species on another as survivability is on stake due to decrease in HSI value. Dependency factor signifies interdependence on each other such as predator-prey relationship. Dependency factor is the ratio of the number of groups of dependent species to total no. of species for a non-zero value of N and 1 for N equal to 0. In extended BBO we have two factors growth rate of species and decline rate of species. We calculate the Extinction rate of species on the basis of emigration rate and decline rate and Evolution rate is calculated on the basis of immigration rate and growth rate[13].

Decline rate = Emigration rate + Extinction rate = $\frac{\mathrm{Ek}\Psi}{\mathrm{n}} + \sqrt{\mathrm{nk}}$ where Ψ is dependency factor, E is emigration rate. (9)

Growth rate= Immigration rate + Evolution rate = $\mathrm{I}(1 - \frac{\mathrm{k}}{\mathrm{n}}) + \frac{\mathrm{c}}{\mathrm{k}}$ where c is constant, I immigration rate. (10)

- k is rank of SIV, n is total number of valid SIV

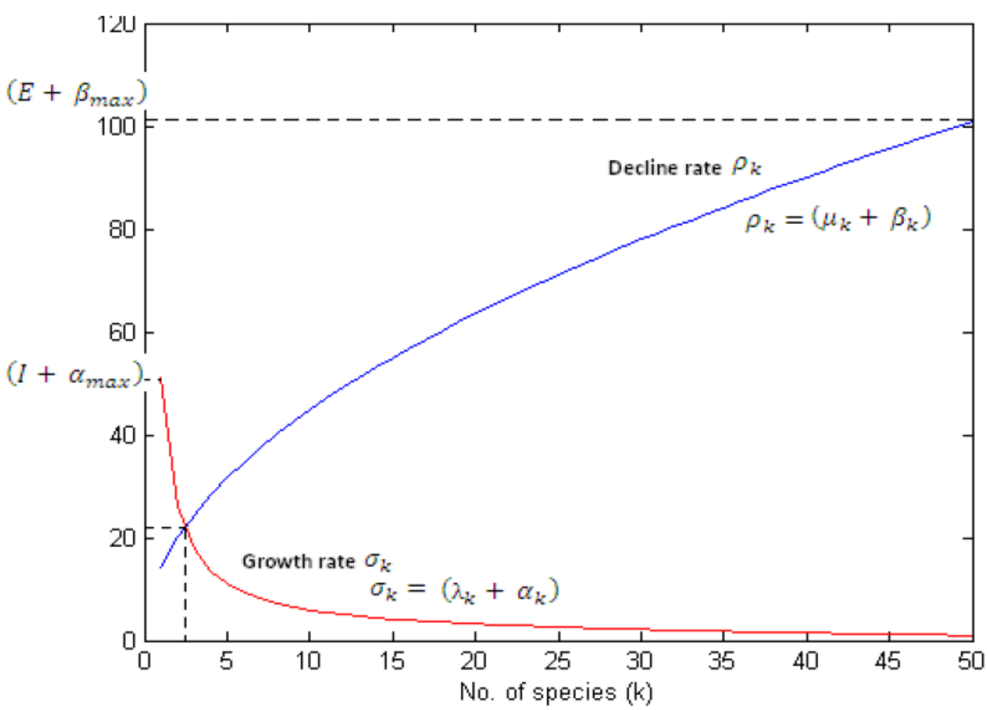

Figure 6 Extended Species Abundance Model

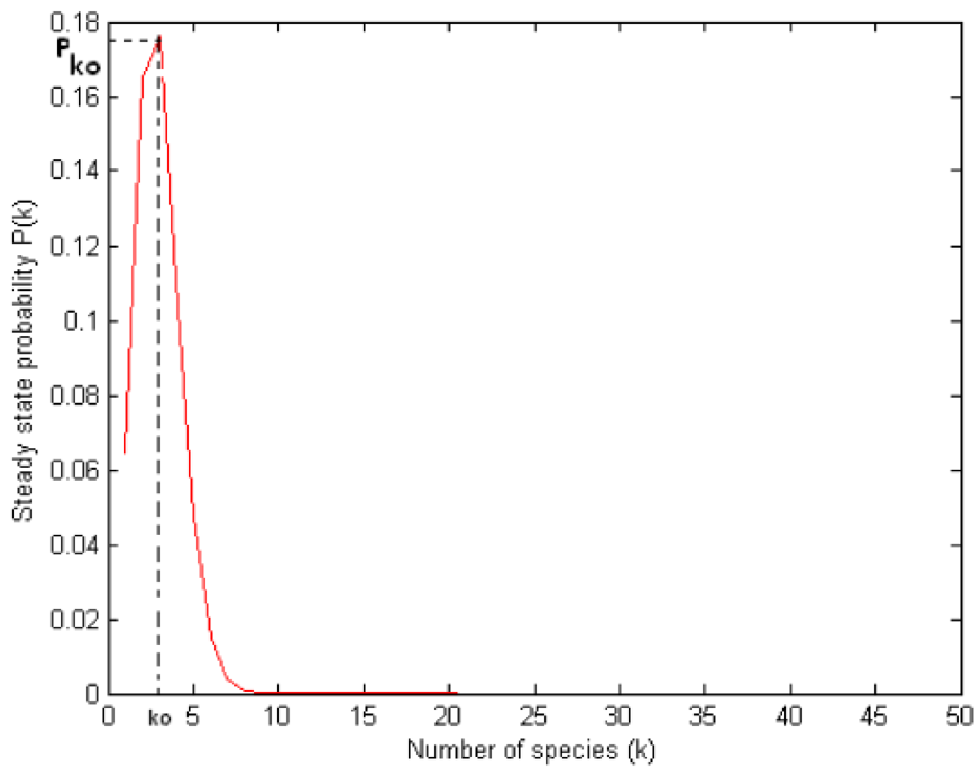

Figure 7 Steady state probability curve P_k as function of species k, k_0 is equilibrium number of species and P_{ko} is that probability

As we can see extended species abundance model as well as probability curve is same in all linear and non-linear migration models[31]. This is because extinction and evolution curves have higher effect than migration curves. Evolution curve is basically a rectangular hyperbola, which overshadows the effect of species growth by immigration and extinction curve which is y centered parabola scales down the effect of species decline by emigration. So this can be concluded that extinction and evolution rate have major role in determining species growth and decline rate rather than immigration and emigration rates. Moreover addition of dependency in determining emigration and immigration decreases their impact in calculating species count in a habitat. Therefore new definition to find growth rate and decline rate is discussed above. In abundance model presented above, maximum value of immigration and emigration are I and E respectively. K_0 is equilibrium point where emigration and immigration rates are equal.

Chapter 4

Computational Intelligence in Face Recognition

In this Chapter we will adapt various evolutionary algorithms like PSO, ACO, Hybrid ACO/PSO, BBO, Extended BBO in face recognition. We will also explain how these techniques are used in the process of face recognition and gives the algorithm for each for recognizing a face.

4.1 Face Recognition Model

Face recognition process is divided into two parts training phase and the recognition phase. In training phase we have feature extraction and feature selection. In our face recognition model we have extracted the features using Gabor kernel and PCA combination and used some evolutionary technique for optimal feature selection. In the recognition phase also, some evolutionary technique is used to carry out the recognition process efficiently.

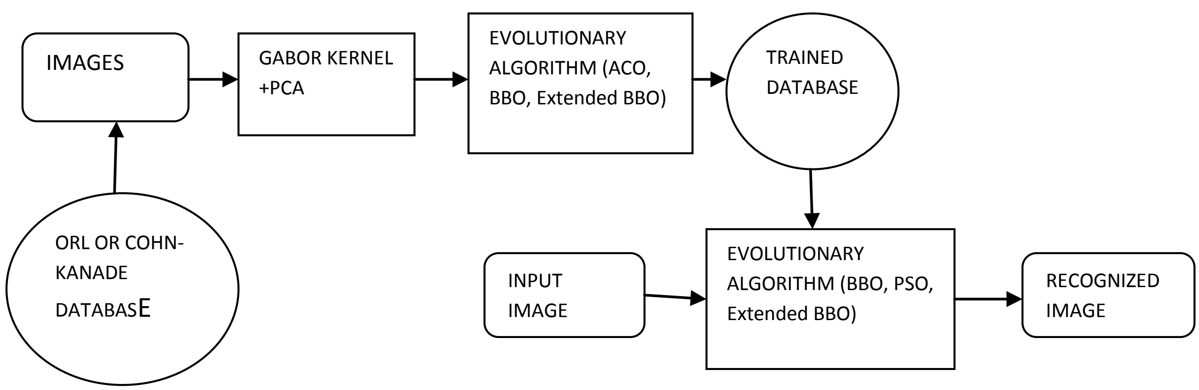

Figure 8 Model for Face recognition

Model of face recognition

1) We take ORL database or Cohn-Kanade database for Face Recognition.
2) We input a set of images for training, which we pass into Gabor kernel to filter and align those images.
3) After proper alignment of images they are input to Principal Component Analysis for features extraction and reduction of dimensions.

4) Features are selected through one of the evolutionary algorithms like ACO, BBO, Extended BBO ,resulted into an optimized and efficient database.
5) Now a new test image is input and that image is recognized through one of the evolutionary algorithms like PSO, BBO, Extended BBO.

4.2 PSO in Face Recognition

In this algorithm in training phase we have two steps. In the first step we filter images with gabor kernel to align images and smoothen them.In the second step we applied PCA to extract features from the training data. In the recognition phase we used particle swarm optimization to reduce the time consumed for identification of a test image.

Various adaptations for PSO[33] in Face Recognition are as follows

Item	**Description**
Particle	An image on the database.
Location/ position	The collection of images (database).
Fitness function	Euclidean distance to measure similarity between input images contained inside the database and input images.
Lbest_position	It is the best position or localization achieved by the particle so far. It corresponds to one of the image
Gbest_position	It is the best position achieved by any particle of the swarm so far
Vmax	Maximum speed of each image used for the comparisons in the complete recognition process.

Table 2 Main elements for PSO in Face Recognition

PSO algorithm for face recognition

1) In the training phase we take a set of images in input known as training set after they are filtered from gabor kernel. We use principal component analysis to extract the features.
2) We represent each image of N*N in form of a vector. Suppose we have M images in training set then a matrix of N^2*M represent all the images in form of vectors.
3) Eigen values and Eigen vectors corresponding to a representation containing all the images in form of vectors is calculated. Now we select k eigen vectors corresponding to k higher eigen values.
4) Now we calculate the weights matrix for each image using the eigen vectors and the initial image vector. The weight matrix of each image represent the proportion of that image in eigen vectors.
5) At the recognition phase, an image to be recognized is taken in input. It is also filtered by gabor kernel and then its weight matrix is calculated using the eigen vector matrix obtained in training phase
6) Weight matrix of all the training images and the weight matrix of the input test image are the two inputs to the PSO algorithm.
7) PSO algorithm is used here in the recognition phase.. Fitness value of each particle is calculated using the fitness function. Here the fitness function is the Euclidean distance between the two images.
8) So fitness is calculated between the image as a current position of particle and the image taken in input for recognition .The weights of both the images are used as the inputs to fitness function.
9) We calculate the fitness of the particle and update the local best position of the particle if its current fitness is better than local best fitness. Now after calculating the fitness of all the particles,the particle which is having the minimum fitness, its current position is taken as the global best position and its fitness value as global best fitness value. This position of particle is basically a image.
10) The current position and velocity of each particle is updated after the first iteration on the basis of local best position and global best position according to the equations (4) and (5).
11) PSO is iterated until we reach maximum count or Euclidean distance between classifier and test image is reached to threshold and after it the value of global best position gives the

image that matches the most with the input test image. We can also tell the class to with the input image belongs using the results of global best position

4.3 ACO in Face Recognition

In training phase we have three steps. In the first step we filter images with gabor kernel to align images and smoothen them. In the second step we applied PCA to extract features and in the last step we use ant colony optimization for feature selection to get a optimial feature vector set.

In the recognition phase we simply use minimum distance classifier to find out the correct class for the input test image.

Ant colony optimization is a meta-heuristic technique for solving computational problems that can reduce the problem into graphs to find good paths. We use ACO in face recognition process to select the optimal set of features[16, 18] . The problem is represented in form a digraph with only 2n or *O*(*n*) arcs instead of a complete graph with O(n2) edges which is generally used in ACO. This modification is done because in the feature selection problem, one feature to be selected is independent with the last feature added to the partial solution. Figure 9 shows how we map our problem in form of a digraph where the nodes represent features of an image, and the arcs connecting two adjacent nodes indicating the choice of the next feature. Given a feature set of size *n*, the feature selection problem is to find a minimal feature subset of size *s* (*s*< *n*) while maintaining a fairly high classification accuracy in representing the original features[].

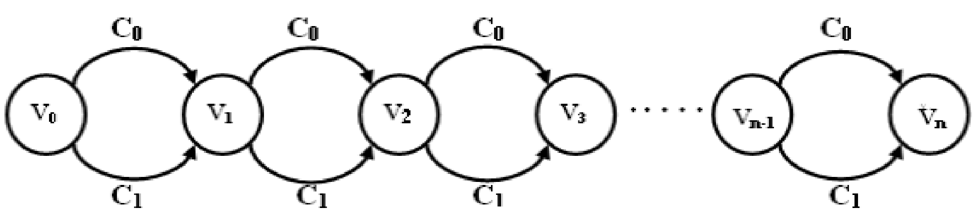

Figure 9 The digraph

Various adaptations for ACO in Face Recognition are as follows

C_i^j	Path from vi-1 to vi for i=1,2…..n and j=0,1

η_i^j	Heuristic value on the path Cij. If j=1, $\eta_i^1 = \frac{\sum_{k=1}^{m}\left(\bar{x}_i^k - \bar{x}_i\right)}{\sum_{k=1}^{m}\left(\frac{1}{N_i^k - 1}\sum_{j=1}^{N_i^k}\left(x_{ij}^k - \bar{x}_i^k\right)^2\right)} \quad (i = 1, \ldots, n)$ Where m is the number of classes in the dataset; n is the number of features; N_i^k is the number of training samples for feature f_i in class k,(k=1……m,i=1,2,….n); x_{ij}^k is the j_{th} training sample for the feature f_i of the images in class k,(j=1,2,….N_i^k); $\bar{x}_i$ is the mean value of the feature f_i of all images;xi-k is the mean of the feature f_i of the images in class k. If j=0 $\eta_i^0 = \frac{\xi}{n}\sum_{i=1}^{n}\eta_i^1$ Where $\xi \epsilon (0,1)$ is a constant
Number of ants	10
Number of iterations	30
Efficiency of the solution	f(s)
τ_i^j	Pheromone on the path Cij. Inintially all trails are initialized by1.

Table 3: Parameters of ACO for face recognition

ACO algorithm for face recognition

1) At the training phase after extracting n features using PCA we apply ACO to optimally select the features. The n features of an image are given as f_1,f_2,….f_n.In graph the i_{th} node v_i is used to represent feature fi. An additional node v_0 is placed at the beginning of the graph from where each ant starts its search[].
2) The two arcs C_i^0 and C_i^1 linking two adjacent nodes v_{i-1} and v_i. If ant at v_{i-1} chooses C_i^j to move to the node v_i then it denotes the desirability to choose feature f_i. If C_i^0 is choosed f_i is selected else not selected

3) We find the solution set containing subset of features as the m ants traces the entire graph.The probability of choosing the path C_i^j(j=0,1) depends upon the pheromone value and the heuristic value of the path.α and β are two parameter that determine the relative importance of the pheromone and the heuristic information.

$$p_i^j(t) = \frac{[\tau_i^j(t)]^\alpha (\eta_i^j)^\beta}{[\tau_i^0(t)]^\alpha (\eta_i^0)^\beta + [\tau_i^1(t)]^\alpha (\eta_i^1)^\beta} \quad (i=1,2,\ldots,n; \quad j=0,1) \tag{11}$$

4) We find the efficiency of all m solution given by m ants and retain the best solution with us. The efficiency of a solution s is given as

$$f(s) = \frac{N_{corr}}{1 + \lambda N_{feat}} \tag{12}$$

where *Ncorr* the number of examples that are correctly classified, *Nfeat* is the number of features selected in *s*, λ is a constant to adjust the importance of the accuracy and the number of features selected.

5) Update pheromons on path traversed by best ant.(Local updation) . Update pheromons of every path(evaporation). When all the ants have completed a solution,the trails are updated by

$$\tau_i^j(t+1) = \rho \cdot \tau_i^j(t) + \Delta\tau_i^j(t) + Q_i^j(t) \tag{13}$$

where

$$\Delta\tau_i^j(t) = \frac{1}{|S_i^j|} \sum_{s \in S_i^j} f(s)$$

$$Q_i^j(t) = \begin{cases} Q & c_i^j \in s_{best} \\ 0 & otherwise \end{cases}$$

S_i^j is the set of solutions generated at the *t*-th iteration passing through c_i^j . S_{best} is the best solution found so far and Q is a positive constant .

6) Iterate step 3 to step5 for best results.

7) Output the best solution matrix after finding best solutions for all the images.
8) The best solution matrix has values in form of zeros and ones. The feature that is selected has value one otherwise zero.
9) At the recognition phase we input the test image and extract the feature using PCA and find the optimal feature subset using ACO. Now compare the input image with optimal features (found in the training phase) of the images in the data set using Euclidean distance .Suppose A and B are two vectors of dimension k.

$$\|A\text{-}B\| = \text{ Euclidean distance} = \sqrt{\sum_{i=1}^{k}(\mathrm{A}^{\mathrm{i}} - \mathrm{B}^{\mathrm{i}})^2}$$

4.4 Hybrid ACO/PSO in Face Recognition

In the training phase we have three steps .In the first step we filter images with gabor kernel to align images and smoothen them. In the second step we applied PCA to extract features and in the last step we use ant colony optimization for feature selection to get a optimial feature vector set.

In the recognition phase we used particle swarm optimization to reduce the time consued for identification of a input test image.

Hybrid ACO/PSO algorithm for Face Recognition

1) At the training phase we take a set of images known as training set. We filter the images using gabor kernel and then features are extracted using PCA. The weight matrix of all the images is obtained. These are the features extracted out of all the features.
2) Now we apply ACO to select the optimal set of features out of the extracted ones. This selection is based on the heuristic value , efficiency etc. The entire algorithm to for feature selection using ACO is given in the last section 4.3.
3) At the recognition phase, we input a test image which is to be recognized. It is filtered by gabor kernel and weight matrix is found.
4) Now we use PSO in the recognition phase. Fitness value of each particle is calculated using the fitness function. Here the fitness function is the Euclidean distance between the two image. So fitness is calculated between the image as a current position of particle and the image taken in input for recognition .The optimally selected weights of the training images

and the weight matrix of the input test image are used as the inputs to fitness function.

5) We calculate the fitness of the particle and update the local best position of the particle if its current fitness is better than local best fitness. Now after calculating the fitness of all the particles,the particle which is having the minimum fitness, its current position is taken as the global best position and its fitness value as global best fitness value. This position of particle is basically a image.
6) The current position and velocity of each particle is updated after the first iteration on the basis of local best position and global best position according to the equations (4) and (5)
7) PSO is iterated until we reach maximum count or Euclidean distance between classifier and test image is reached to threshold and after it the value of global best position gives the image that matches the most with the input test image. While finding the Euclidean distance we find the distance between the optimal set of features and the image having the minimum sum is the best match.We can also tell the class to with the input image belongs using the results of global best position.

4.5 Biogeography Based Optimization(BBO) in Face Recognition

In the training phase we have three steps In the first step we filter images with gabor kernel to align images and smoothen them. In the second step we applied PCA to extract features and in the last step we use biogeography based optimization for feature selection to get a optimial feature vector set[8].

In the recognition phase we have used biogeography based optimization to recognize an input test image. There are various parameters in BBO used at training and recognition level.

Suitability Index Variable	Each feature in an image act as SIV. And each image act as habitat. In ORL database 119 SIVs are used as 120 training images and in Cohn-Kanade database 79 SIVs are used as 80 images are used for training.
Habitat Suitability Index and Fitness Value	Euclidean distance acts as fitness value. Accuracy of a habitat is predicted through HSI. Migration is done to achieve high HSI.

Emigration Rate	Emigration is leaving one's own native place .Emigration rate = $\frac{Ek}{n}$ where E is emigration rate, k is rank of SIV, n is total number of valid SIV. E is taken as 1.
Immigration Rate	Immigration is arrival of a species to new place. Immigration rate = $I(1 - \frac{k}{n})$ where I immigration rate and taken as 1.

Table 4 Parameters of BBO for face recognition

BBO algorithm for Face Recognition

1) A database is taken into input and divided into 2 parts: train data and test data. Both types of data passes through the gabor kernel filter and all the images get aligned and smoothened. Now we apply PCA on train and test data. We get the extracted features from the train and test data named train features and test features respectively.
2) Now on applying BBO we try to train the test features using the train features. This is don as follows:

- HSI of the train features is calculated through fitness function(standard deviation) and is termed as ideal HSI.
- Now we iterate BBO= number of SIVs and in each iteration we do the following:

➢ Find the HSI value for the SIV(or column) of the image in the test features matrix

➢ Compare the calculated HSI with the ideal one. If it is close to ideal one do nothing and if not perform migration to improve the test features.

We now calculate similarity matrix from test and train features using Euclidean distance which shows how better test features are trained. We use these test features in the recognition phase.

3) In the recognition phase, we input a image, its features are extracted using Gabor kernel and PCA. Now we try to recognize this image using the test features obtained from training phase and the features of this image. In solution we try to find out an image whose maximum features are close to the features of input image.
4) The effectiveness of this algorithm is that inspite of matching the image with whole database we compare the image only with probable solutions.
5) We select k SIVs randomly. Now we find the m images which are closest to the input image

on the basis of these k SIVs.

6) We add these m images in probable solution and iterate this for a number times and union the images in the probable solution in each iteration.
7) Emigration Rate of each SIV is calculated and they are emigrated according to decreasing order of their emigration rates.
8) Immigration rate corresponding to the emigrated SIV are calculated and they are immigrated.
9) Euclidean distance between the images in the probable solution and the image in the input is found. The images having the minimum distance termed as ideal solution.
10) Mode of all images of ideal solution are calculated and distance is calculated between recognizing image. If it is less than threshold process stop or if maximum iterations are reached.

4.6 Extended BBO in Face Recognition

In the training phase we have three steps In the first step we filter images with gabor kernel to align images and smoothen them. In the second step we applied PCA to extract features and in the last step we use BBO based on extended species model of Biogeography for feature selection to get a optimial feature vector set.

At recognizing phase we have used BBO based on extended species model of Biogeography to recognize an image. In this we input an image and apply algorithm to find a best match without comparing our input image with all the images in training database. This led to reduction in time and makes it an efficient approach.

For adapting extended BBO in face recognition application we have to set various parameters, they are stated in table 5:

Suitability Index Variable (SIV)	Each feature in an image act as SIV. And each image act as habitat.
Number of SIVs	119 for ORL database and 79 for Cohn-Kanade database
Number of iterations for Extended BBO algorithm	Number of SIVs or threshold is reached
Habitat Suitability Index	Euclidean distance acts as fitness value. Accuracy of a habitat is

and Fitness Value	predicted through HSI. Migration is done to achieve high HSI.
Extinction Rate	Extinction rate = $\frac{Ek\Psi}{n}+\sqrt{nk}$ where Ψ is dependency factor, E is emigration rate, k is rank of SIV, n is total number of valid SIV
Evolution Rate	Evolution rate=$I(1-\frac{k}{n})+\frac{c}{k}$ where c is constant, I immigration rate ,
Dependency Factor	At training phase correlation between SIV is checked to measure dependency to migrate SIV from one habitat to other. And at recognizing phase dependency is to find probable solutions when one habitat of particular subject is there, remaining other habitat are also used as solutions.
Factor Of relevance	At training phase each SIV has a particular weight so, If that weight is lower than a threshold it is eliminated from process of evolving. At recognizing phase probable solutions are find according to relevance of feature, solution of closer distance to higher relevant SIV are taken as probable solution to fasten up recognizing process .

Table 5: parameters of extended BBO for face recognition

Extended BBO algorithm for Face Recognition

The algorithm is divided into training phase and recognition phase.

Training Phase:

1) We take the dataset in input and divide it into 3 parts i.e training set, test set and eval set. We extract features from all the sets using PCA and Gabor kernel combination and they are termed as train features, test features and eval features respectively. We use the train features and the eval features in the training phase. Train features are used to train the eval features.
2) Compute a matrix (eval_matrix) that is a average of set of images of each subject of eval features,so that we have a solution of highest HSI. Then we will generate multiple solutions to optimize database using extended BBO.
3) Now in eval matrix we check for relevance factor. If weight of the SIV is less than a threshold value then do nothing else calculate the extinction rates of SIVs and sort them in decreasing order according to their extinction rates.

4) Calculate the correlation factor of the SIV with highest extinction rate with all other SIVs. If it comes more than 0.7 then don't migrate it. In that case choose the SIV with next highest extinction rate from step 3.

5) Calculate growth rate of SIV for all solutions corresponding to extinction SIV. Migrate highest growth rate SIV to eval_matrix. If new SIV value minus old SIV value is less than threshold than stop this process else go to step 3 with next higher extinction rate .

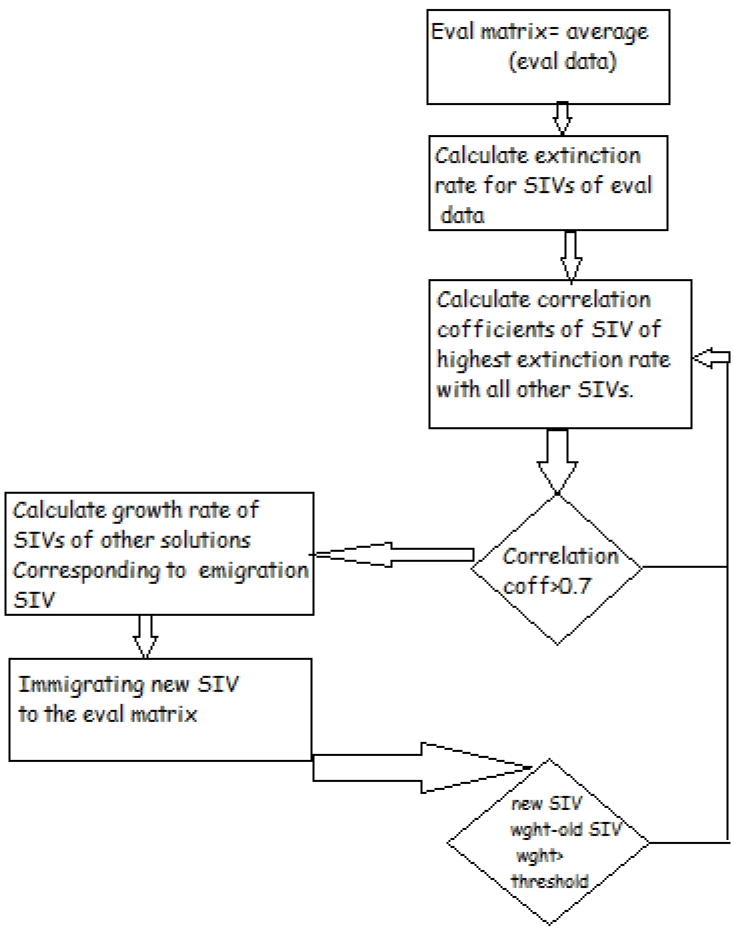

Figure 10 Architectural Flow diagram for training algorithm of Extended BBO

6) Repeat for all the classes of database.

7) Calculate similarity matrix from train_data and test_data using Euclidean distance and plot the required ROC curve.

Recognition phase :

1) An input image is taken for recognition. We will extract features using Gabor kernel and PCA and then image is recognized using Extended BBO.
2) We calculate the SIVs of higher relevance and then we find closest images according to these SIVs are . These images constitutes the probable solution.
3) Extinction rates are calculated of each SIVs and SIVs are migrated according to their extinction rates. On the contrary in BBO SIVs were migrated according to the immigration rates.

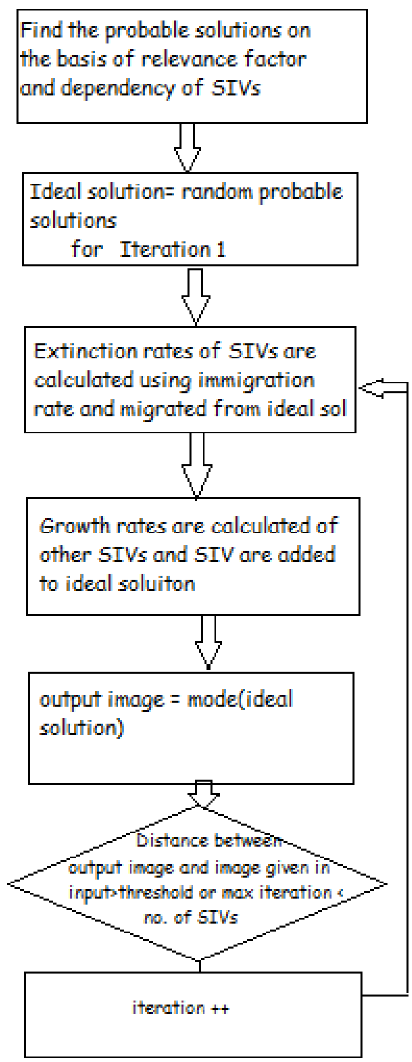

Figure 11 Architectural Flow diagram for recognizing algorithm of Extended BBO

4) Now according to growth rates , SIVs are migrated into the ideal solution along with image id increasing the HSI of the ideal solution.
5) Mode of all the images of ideal solution is calculated. Distance is calculated between it and recognizing image, if it is less than a threshold value process stops or continued up to number of SIVs iterations

This chapter explained various evolutionary techniques in the area of Face Recognition. Now in the next chapter we are going to compare all these techniques on the basis of running time and accuracy and analyze which one would be best to use.

Chapter 5

Experimentation, Evaluation and Results

In this chapter we will compare all the evolutionary techniques explained in previous chapter on the basis of recognition time and accuracy. We have tested and shown our results on two datasets.

5.1 Databases Taken For Experimentation

We have taken two databases for experimentation. For simulation purpose we have used standard ORL database figure 12 which is acquired at the Olivetti Research Laboratory in Cambridge, U.K and Cohn-kanade AU- coded facial expression by Kanade, T.Cohn, J.F., & Tian, Y.(2000). ORL database consist of 400 different images of 40 subject i.e. 10 images of each subject. Each image is a gray level image with size of 92*112 pixels [40]. Each image has different characteristic like illumination, facial and pose expression.

Figure 12 ORL Database

Figure 13 Cohn-Kanade Database

Cohn Kanade database as shown in fig13 we used consists of 240 images of 8 subjects i.e. 30 images of each subject. Each image is a grey level image with size of 640*490 pixels.

5.2 Experimental Setup

For our experiment, we used both ORL database and Cohn-Kanade database**.**

Training Phase

- In case of Exhaustive, PSO, Hybrid ACO/PSO and exhaustive search : On using ORL database we used 360 images for training 9 from each class as there are 40 classes. On using Cohn-Kanade database we used 240 images for training 30 from each class asthere are 8 classes.

  But in case of ACO, while extracting features using PCA at time time of selecting k eigen vectors of high k eigen values we take the value of k=15. So, 15 are the number of features extracted from each image after PCA. In case of ORL database ACO optimally selects 12 features out of 15 and in case of Cohn-Kanade database it selects 9 features out of 15 for each image. ACO reduces the number of features which helps in reducing the recognition time

- In case of BBO: We divide the training database into two parts, train data and test data. On using ORL database ,each train data and test data have 120 images,3 from each class. So number of SIVs in this case are 119. On using Cohn-Kanade database each train and test data has 80 images,10 from each class.So number of SIV in this case are 79.

- In case of Extended BBO: We divided our complete database into 3 parts: test data, eval data and train data. In ORL database at training time we used 360 images. We have 40 subjects so 9 images correspond to each of the subject .Each set of images contains 120 images i.e. a 119 *120 matrix [10].

There are exactly 119 SIVs, corresponding to each of the most efficient features *n* is the number of SIVs which have weight more than 1, *k* is the rank of SIV, *c* is also taken as 1 and the correlation coefficient is taken as 0.7. In Cohn-Kanade database at training time 240 images are used and divided into 3 parts : eval data, train data and test data. Each set of images contains 80 images i.e a 79*80 matrix. There are exactly 79 SIVs.

We train our system through ideal_data and train_data. Using the results we developed a train model. We used Test_data to test our train model and we showed our results in the form of ROC plot.

Recognition phase :

In the recognizing phase trained database is used to recognize the image. A simulation is done on 120 images of ORL database to test the accuracy.A simulation is done on 160 images of Cohn-Kanade database to test the accuracy.

We are experimenting on two different kind of databases. ORL database has more number of classes but less number of training samples in each class and on the other hand, Cohn-Kanade database has less number of classes but more number of training samples in each class.

5.3 Results And Analysis

In this section we are going to compare various algorithms on the basis of results.

For the comparision in training phase ROC curve is shown of BBO and Extended BBO. We have drawn ROC curves to shown how well the training data is trained .

Performance of Training database of our face recognition based on extended BBO is calculated by finding a similarity matrix, which consist of comparisons of each image with all the remaining images and false acceptance rate (FAR) and false rejection rate (FRR) are evaluated.

$$FRR = (a/b) * 100\% \tag{14}$$

Where *a* = number of rejected genuine identity claims and *b* = total number of genuine claims made

$$FAR = (c/d) * 100\%. \tag{15}$$

Where *c* = number of accepted false identity claims and *d* = number of false identity claims made. And verification rate is equal to *1-FRR*. ROC curve between extended BBO and PCA+ Gabor is shown in below fig.14.

.

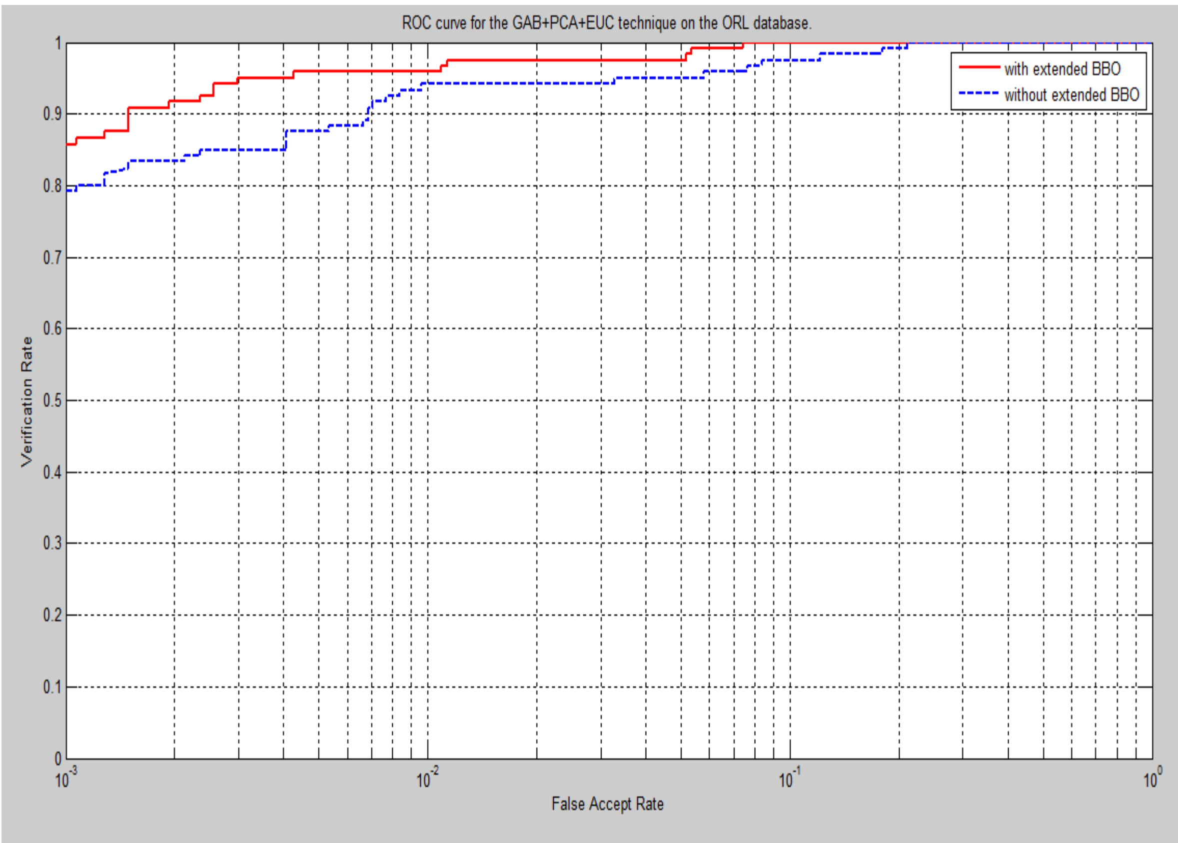

Figure 14 ROC Curve between Extended BBO and without it

And ROC curve between extended BBO and BBO is shown in fig.15. It can be clearly seen extended BBO outperform PCA and Gabor and BBO

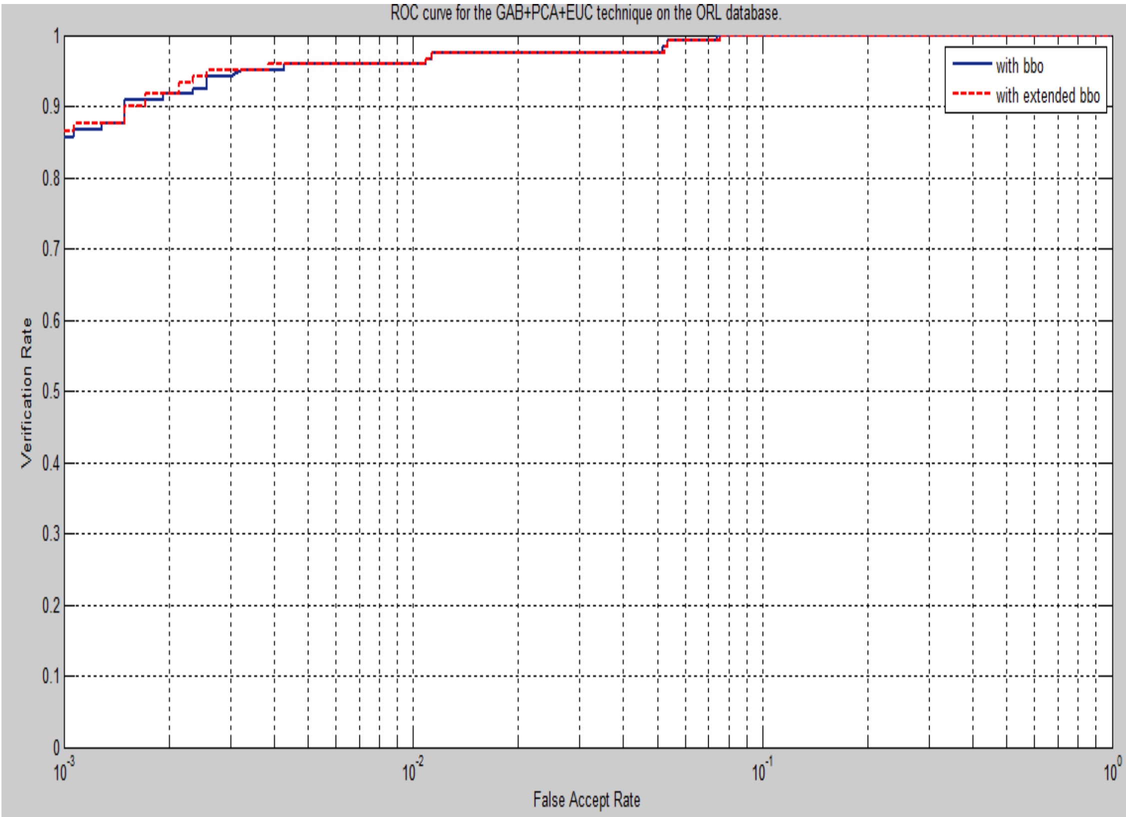

Figure 15 ROC Curve between Extended BBO and BBO

In the recognition phase we compare the algorithms on the basis of accuracy and time taken for recognition.

On the basis of time taken for recognition:

- We have analysed that time taken by PSO and BBO is minimum in comparision to other algorithms. Though PSO takes least time but it has spikes which even crosses the time of exhaustive search.
- ACO takes the maximum time for recognition. The reason behind this is in this algorithm we use ACO in training phase and minimum distance classifier(matching with all the images) in recognition phase. Since matching is done with all the images it takes the most time and additional time is added of using ACO in feature selection. So, ACO takes more than exhaustive search and Extended BBO.

- In Hybrid ACO/PSO, instead of minimum distance classifier in the recognition phase we use PSO which reduces the time as PSO compares with probable random solutions and not with all the images.

	Exhaustive Search	ACO	PSO	Hybrid ACO/PSO	BBO	Extended BBO
ORL database	**.009012**	**0.03205**	**0.00526**	**0.0237**	**0.00544**	**.008591**
Cohn-Kanade database	**0.00715**	**0.0340**	**0.00264**	**0.0274**	**0.00407**	**0.00858**

Table 6 Time taken by various algorithms on both databases

- Extended BBO recognizes on the basis of probable solutions generated on the basis of high relevance SIVs and BBO recognizes on the basis of probable solutions on the basis of randomly generated SIVs, so they take less time as one to one matching is not there. Time taken by BBO is less then Exhaustive search and Extended BBO takes time approximately close to Exhaustive search.

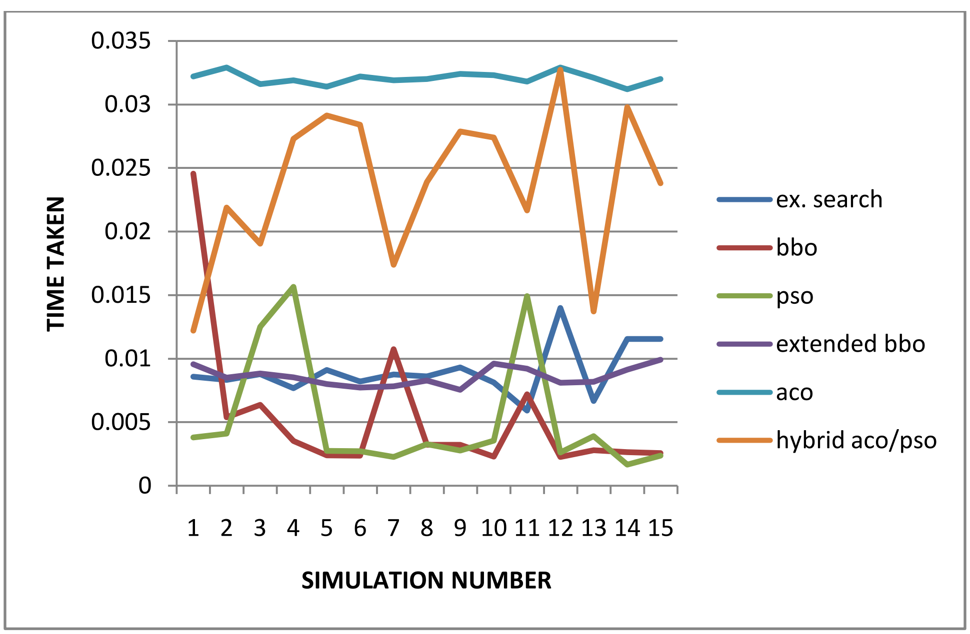

Figure 16 Graph to show simulation results of various algorithms on ORL database

- According to recognition time the sequence of algorithms is going to be

PSO < BBO < exhaustive Search < Extended BBO < hybrid ACO/PSO < ACO

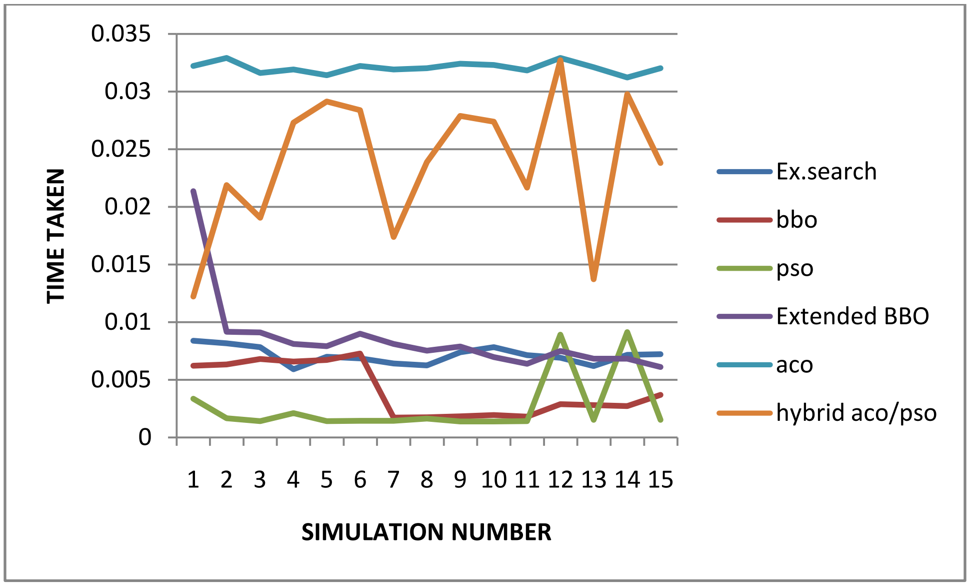

Figure 17 Graph to show simulation results of various algorithms on Cohn-Kanade database.

On the basis of Accuracy:

- Extended BBO has the highest accuracy in comparision to all other algorithms. Its accuracy is approximately equal to the accuracy of exhaustive search.
- Though PSO takes least time for recognition but it has least accuracy.
- ACO has higher accuracy than PSO because in recognition phase it uses minimum distance classifier as fitness function which is matching with all the images.
- According to accuracy the sequence of algorithms is going to be

Extended BBO < Exhaustive Search < BBO < ACO < Hybrid ACO/PSO < PSO

	Exhaustive Search	**ACO**	**PSO**	**Hybrid ACO/PSO**	**BBO**	**Extended BBO**
ORL-database	**90.58**	**83.33**	**50.52**	**81.63**	**84.78**	**92.5**
Cohn-Kanade database	**96.55**	**87.11**	**68.23**	**85.93**	**85**	**98.27**

Table 7 Accuracy in both the databases

Now on seeing both the sequences it depends on us which parameter is more important for us. Both if we want to find an appropriate algorithm Extended BBO has an high accuracy with fairly minimum amount of time.

.

Chapter 6

Conclusion and Future Work

In this thesis we compared various nature inspired algorithms. We compared some advance algorithms like BBO and Extended BBO with some elementary evolutionary algorithms like ACO, PSO and Exhaustive Search. We also proposed an hybrid algorithm i.e. hybrid of aco/pso using one evolutionary technique in training phase and other evolutionary technique in recognition phase. We used ACO in training phase for feature selection and PSO in recognition phase to reduce the recognition time using probable solutions for recognition. This technique is also compared to Extended BBO and BBO and has less efficient results than them,both in terms of accuracy and time.

For face recognition, first we have to extract features from the image then develop a training database then recognize the given input image with the help of training database developed in training phase. For feature extraction we have used PCA in all the algorithms. PCA has one drawback, it cannot take raw images directly as input as it requires properly aligned and constantly illuminated images. So for this we have used gabor kernel which smoothen images and illuminate them evenly and set them in proper alignment. After this we pass our images in PCA algorithm which results in extraction of features.

After feature extraction we select features using various evolutionary techniques. In our thesis we have used ACO, BBO, Extended BBO for this phase. In the recognition phase we use different techniques like PSO,BBO, Extended BBO. We can also use different evolutionary algorithm in training phase and different in recognition phase. We have developed one algorithm on this basis which is Hybrid ACO/PSO.

We compared all these algorithms on the basis of running time for recognition and accuracy and we have analysed that different algorithm is better for different parameter. If we consider time, PSO is best as it takes least time but its accuracy is worst.

On the basis of time we have :

PSO < BBO < exhaustive Search < Extended BBO < hybrid ACO/PSO < ACO

On the basis of accuracy we have :

Extended BBO < Exhaustive Search < BBO < ACO < Hybrid ACO/PSO < PSO

Both if we want to find an appropriate algorithm Extended BBO can be our choice. As its running time is almost equal to exhaustive search or sometimes less than it but it has an high accuracy.

6.1 Future Work:

For feature extraction phase we have used Principal Component Analysis. We can use some otherfeature extraction technique here for example Discrete Cosine Transformation, Linear Discriminant Analysis and then we can apply some other evolutionary technique for feature selection and some technique for recognition.

We are using evolutionary algorithms for feature selection and recognition. Recognition phase is basically classification. We can also use support vector machines or minimal complexity machine for classification.

References

[1]Ahmed, N., Natarajan, T., & Rao, K. R. (1974). Discrete cosine transform. *IEEE Transactions on Computers.*

[2] Wein-Lun Chao,"Face Recognition Survey", *GICE, National Taiwan University*,2010

[3] Bhuiyan, A. A.; Liu, C. H. (2007). On face recognition using gabor filters",. *In Proceedings of World Academy of Science, Engineering and Technology*.

[4] D., C., He, X., & Han, J. :. (2007). Semi-supervised discriminant analysis" Computer Vision. *IEEE 11th International Conference on Computer Vision.*

[5] Brunelli, R., & Poggio, T. (1993). Face recognition: features versus templates. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 15.*

[6] Chung, K.-C., Kee, S. C., & Kim, S. R. (1999). Face Recognition Using Principal Component Analysis Of Gabor Filter Responses. *IEEE*, (pp. 53-57).

[7] Struc, V., & Pavesic, N. (2009). Gabor-Based Kernel Partial-Least-Sqares Discrimination Features For Face Recognition. *Informatica*, (pp. 115-138).

[8] Gupta, D., Goel, L., & Abhishek. (2013). An Efficient Biogeography Based Face Recognition System. cse-13.

[9] Turk, M. (2001). A random walk through eigenspace. *IEICE Transactions on Information and Systems.*

[10] Turk, M., & Petland, A. (1991). Eigenfaces for recognition. *Journal of cognitive neuroscience.*

[11]Dorigo, M. (2007). Ant Colony Optimization. *Scholarpedia,*.

[12]Fisher, R. A. (1936). The use of multiple measures in taxonomic problems. *Annals of Eugenics*, *7.*

[13]Goel, L., & Gupta, D. (2012). Extended Species Abundance Models of Biogeography Based Optimization. *Fourth International Conference on Computational Intelligence, Modelling and Simulation.* IEEE.

[14]Goel, L., Gupta, D., & Panchal, K. V. (2011). Hybrid Bio-Inspired Techniques For Land Cover Feature Extraction. *Elsevier.*

[15]Goel, L., Gupta, D., & Panchal, V. K. (2012). Dynamic Model of blended bigeograhy based

optimization for land cover feature extraction. *Springer CCIS 306.*

[16] Ling Chen1,2, Bolun Chen1, Yixin Chen3 , "Image Feature Selection Based on Ant Colony Optimization" in *Springer* ,2011.

[17] Mehdi Hosseinzadeh Aghdam *, Nasser Ghasem-Aghaee, Mohammad Ehsan Basiri, “Text feature selection using ant colony optimization” , in *Elsevier*, 2007.

[18] Amrita1,Nirvair Neeru2, “A Novel Approach for Medical Image Stitching Using Ant Colony Optimization”, *1Student of M. Tech Computer Science, 2Assistant professor, Punjabi University, Patiala*, May 2014.

[19] Esra Saraç and Selma AyGe Özel, “An Ant Colony Optimization Based Feature Selection for Web Page Classification” , *Department of Computer Engineering, C¸ukurova University, Balcali, Sarıc¸am, 01330 Adana, Turkey,* July 2014.

[20]Holden, N., & Frietas, A. A. A Hybrid PSO/ACO Algorithm for Classification.

[21]Howell, A., & Buxton, H. (1996). Towards unconstrained face recognition from image sequences. *Proceedings of the Second IEEE International Conference on Automatic Face and Gesture Recognition.*

[22]Jafri, R., & Arabnia, R. H. (2009). A Survey Of Face Recognition Techniques. *Journal of Information Processing Systems.*

[23]Jafri, R., & Arbnia, H. R. (2009). A Survey of Face Recognition Techniques . *5*.

[24]Jollife, I. T. Principal Component Analysis.

[25] WANG Lu, WANG Zhiliang, HU Siquan, LIU Lei, “Ant Colony Optimization for Task Allocation in Multi-Agent Systems” , *School of Information Engineering, University of Science and Technology Beijing, Beijing 100083, China*.

[26]Kanan, H. R., Faez, K., & Hosseinzadeh, M. (2007). Face Recognition System Using Ant Colony Optimization-Based Selected Features. *IEEE Symposium on Computational Intelligence in Security and Defense Applications.*

[27]Kaur, N., Gupta, P., & Kaur, A. “Face recognition using biogeography based optimization.

[28]Kennedy, J., & Eberhart, R. (1995). Particle swarm optimization. *IEEE International Conference on Neural Networks.*

[29] Simon, D. (2008). Biogeography Based Optimization. *IEEE.*

[30] Murugan, D., Arumugam, D. S., Rajalakshmi, K., & Manish. (2010). Performance Evaluation of face recognition using Gabor filter, log Gabor filter and discrete wavelet transform. *International jouranal of computer science & information Technology*.

[31]Ma, H. (2010). An analysis of the equilibrium of migration models for biogeography-based optimization. *nformation Sciences 180.*

[32] Turk., M., & Pentland, A. (1991). Face Recognition Using Eigenfaces. *IEEE Conference on Computer Vision and Pattern Recognition.* IEEE.

[33] Villegas, O. O., & Quintero, M. A. (2009). A Novel Evolutionary Face Recognition Algorithm Using Particle Swarm Optimization. *Fifth International Conference on Signal Image Technology and Internet Based Systems.*

Appendix –A

Abbreviations

PCA – Principal Component Analysis

BBO – Biogeography Based Optimization

PSO – Particle Swarm Optimization

ACO – Ant Colony Optimization

PC – Principal Component

ROC – Receiver Optimizing Characteristic

ORL – Olivetti Research Laboratory

LDA – Linear Discriminant Analysis

Extended BBO – Extended Species abundance model of biogeography

AS – Ant System

FAR – False Acceptance Rate

FRR – False Rejection Rate

HMM – Hidden Markov Model

HTER – Half Total Error Rate

HSI – Habitat Suitability Index

DE – Differential Evolution

Appendix – B
Matlab

The short form for MATLAB is "matrix laboratory". The software MATLA can be said as the powerful software package which has many functions which are in-built for the fulfillment of large and different types of tasks. Types of task may range from small and simple mathematical operations to the complex 3-dimensional imaging. In addition to all this MATLAB also has wide variety of programming constructs that allows the user to modify as well as customize the programs according to their own requirements, application and specification. The common uses of MATLAB includes exploration and visualization, Data analysis, Algorithm, simulation and prototyping, Scientific and engineering, simulation and prototyping, visualization and Application development, exploration and visualization, Data analysis, Modeling, Math and computation, Algorithm, including graphical user interface building.

MATLAB can be referred as an interactive system. Basic element of the MATLAB can be said to be an array which does not require any sort of dimensioning. So, many of the problems especially related to technical computing (for example those related to the matrix and vector formulations) can be solved in much lesser time or more appropriately in fractions of time. Same problems would have taken much larger time if program was written in non-interactive and scalar languages as that of C or FORTRAN.

The language MATLAB was written originally so that easy access can be provided to the matrix software developed especially by the EISPACK and LINPACK projects. Now a day software's established by ARPACK and LAPACK projects are used by the MATLAB.

one of the feature of MATLAB is that it provides a wide of family of solution to some application specific problem i.e. toolboxes. This feature of MATLAB allows the user to apply and learn the specialized technology. Toolbox can be said to be as collection of the M-files (MATLAB functions) which deal with all or nearly all application of particular requirement.

64

Particular class of problem can be solved with the help of toolbox. Various area in which toolboxes are made available are wavelets, simulations, control system, neural networks, signal processing, and many others.

As MATLAB is typically used as a specification rather than an implementation language, tools to translate MATLAB code to the target programming language are important to achieve highlevels
of productivity and efficiency. However, in order to attain a given desired efficiency level, the high abstraction level provided by MATLAB requires that tools be guided by users. In fact, the issues described above arise often in relation to automatic synthesis of MATLAB specifications to a software language or a hardware description language. In the steps for finding efficient implementations, users have to conduct customized profiling schemes, monitoring techniques, and data type and word length exploration, mostly through the invasive insertion of new code. (Cardoso, Fernandes, Monteiro, Carvalho, & Nobre)
So, in short MATLAB is commercial mathematical software produced by Math Works Inc., which is used to develop mathematical algorithms, data visualization, data analysis and senior technical calculation language and interactive environment of number calculation, mainly including two parts, MATLAB and Simulink. MATLAB not only has plenty of function modules and tool boxes, which can provides the functions such as communication, radio frequency, filtration and wavelet analysis, but also it can have good interface with C/C++ language, and can change M documents written by MATLAB language into C/C++ documents automatically, also support users to do independent application development. MATLAB has the advantages as follows: high efficiency in programming, convenience, high expansion ability and strong portability, simple language, and convenient mapping functionality, efficient matrix and array operation, etc. which can afford a great help for teachers teaching and students learning by themselves. The contents of communication theory course can be realized through interactive GUI interfaces by students' themselves, which not only lets the students have a good understanding of the abstract contents, but also through the realization of programming which can train the students' ability to solve problems and develop students' abilities of programming and innovation. (Guangpua & Yuchun, 2012).