CHAPTER 1

INTRODUCTION

Face recognition has theoretical interest from cognitive scientists as well as practical importance, so it's a real world problem since evolution of computer vision. Face is a common thing to identified and face recognition is non-invasive thus it has always a major area of research.

Kohonen uses a very early face recognition system demonstrating alignment an normalization of facial images with the help of artificial neural net [22]. Kohonen uses eigenfaces for recognizing the images. Eigenfaces are the eigenvectors of autocorrelation matrix of image to be recognized. Although this system was not very successful in practical. Since then different methods like inter face distance, recognition based on edges etc are comes into existence. Some of them were successful when aligned images are used but in case of large databases and images locating the different scale of faces none of them were successful Later an algebraic manipulation that shows less than 100 were required to aligned and normalize facial images and was easy to calculate them, introduced by Kirby and SIrovich [21]. According to Turk and Pentland [19] residual error upon coding using eigenfaces can be used to detect faces in natural imagery and can also be used to determine the scale of faces in images.

With the advancement of technology, it is bound to be fulfilling all the needs of human being. Any new application or development should be effective and easy to be used. The main parameter of an application or a system is security in this era of information. Primary demand of systems is to secure our assets and privacy. Biometrics fulfills this primary demand and examples of biometric system are voice recognition, iris, face etc. Face recognition is the most collectable and accessible among all of biometric systems.

Every human have its own and unique characteristics and these are used in face recognition system. Due to efficient optimization algorithm available in computer vision it becomes the most useful application in security and privacy. Identification and verification are the two main uses of face recognition systems, in former matching is done one by one while in later one to many matching is performed. Face recognition may not be the most efficient and reliable over the other biometric techniques but it has some advantages. This is very useful in case of mass scanning like at railway and airport authority for scanning the suspicious objects.

Many of nature computational intelligence, optimization algorithms are very efficient and can be applied directly to the real world complex problems. Metaheuristic algorithms such as Biogeographical Based Optimization(BBO), Ant Colony Optimization(ACO), Hybrid ACO/PSO, Particle Swarm Optimization(PSO) , Extended BBO and traditional approaches like Principal Component Analysis(PCA) have been used in the field of face recognition. These approaches force the user to be confused to choose the best method for recognition.

1.1 Motivation

Some of the conventional methods use the term “what we have” for authentication such as ID cards, password but the term “who we are” is much better the former. Who we are can be identified using biometric authentication like iris, fingerprint, face recognition. Like other biometric systems Face recognition is non-alterable to a human being.

DCT, PCA, DWT [6] are the various techniques used to feature extraction but PCA is the most common one. Feature based system’s performance depends upon age span, illumination facial pose and several other parameters. Geometric facial expression are used in early system of facial recognition but now a days several technique have been found like nature inspired algorithm and neural networks. Principle component analysis uses the best vectors for distribution of facial images within the space by dimensional reduction. These vectors are called eigenfaces.

D. Simon [11] presented the Biogeographical based optimization algorithm based on the biological organism distribution. This can be applied to the problems with high dimensionality and those have multiple local optima. BBO promotes the uses of biogeography concepts in optimization problems.

Kennedy and Eberhart [18] proposed a evolutionary optimization algorithm named as Particle Swarm Optimization. PSO treat each individual as a particle and composition of particle as a swarm. PSO is a population based optimization technique and behave like a bird flock searching for their food and area. Ant Colony Optimization, proposed by M. Doringo et al [12] is an evolution optimization algorithm simulating the social behavior of ants. Travelling salesman problem, graph coloring system fault detecting are some of the main application of ACO. Computational time of the various problems can be optimized with the use of heuristic and random search strategies by ACO.

BBO performance on a set of 14 standard benchmarks demonstrated by Simmon [11]. Extended BBO uses extended species model of BBO and shows interdependencies on various species over predator prey relationship. Evolving process considered only those states whose particular SIV is greater than some threshold value. BBO is efficiently used in embedded problem and various optimization problems. Extended BBO involves the decline and growth rate as a function of emigration rate, evolution rate and immigration rate [4]. BBO and Extended BBO has been used for Face Recognition [5].

Grey wolf optimization algorithm stated by S. Mirjalili et al [2] is inspired by social behavior of grey wolves. The GWO algorithm mimics the leadership hierarchy and hunting mechanism of grey wolves in nature. Four types of grey wolves such as alpha, beta, delta, and omega are employed for simulating the leadership hierarchy. It has shown good result in selecting feature subsets in data mining, multimedia information system, pattern classification and machine learning system [1]. These results motivated us to apply GWO for Face Recognition.

1.2 Related Work

The Eigenface method is one of the generally used algorithms for face recognition. Mathematically, Eigenfaces are the principal components divide the face into feature vectors.

Covariance matrix is used to recover the feature vector information. These Eigenvectors are used to quantify the variation between multiple faces. The faces are the linear combination of highest Eigenvalues. Some other approaches are Discrete Cosine Transform [28], Linear Discriminant Analysis [24], principle Component Analysis [32].many researchers proposed unsupervised learning at feature extraction while supervised learning is proposed at recognizing phase that helps in reducing the classification errors. Raw images can't be inputted to the algorithms directly to work efficiently so other pre condition like aligning the images , contrasting the illumination is required. Principle Component Analysis is hybridized with Gabor Filter which helps in achieving high accuracy by Chung et al [17].

Recently some other metaheuristic algorithms are hybridized like PSO and ACO and shown efficient result in face recognition. BBO is used with both of Principle Component Analysis and Discrete Cosine Transform [3]. Extension of BBO is recently used by Lavika Goel [3] for edge detection in remote sensing of images.

1.3 Problem Statement

In this era of information system security is the primary concern of user so there is need of authentication and verification. Face recognition system plays an important role authorizing and verifying the user. To match a given input image with other single image or from a set of images from database is the goal of face recognition problem. We need to find an efficient and accurate algorithm for this purpose. Face recognition involves detection of faces and feature extraction and then classification.

There is a database of images and an input image is given; and the goal is to found that whether there is a match between input image and images in the database. In previous approaches PCA was used in feature extraction and various metaheuristic algorithm like ACO, BBO, PSO, extended BBO were used. Grey Wolf optimizer (GWO) has shown promising and more than satisfactory results in the various domains like Thermal power systems [32], feature subset selection [31], time forecasting [33], vehicle routing [34] and optimizing key values in cryptographic algorithms [35].

Focus of our research will be on the exploration of GWO algorithm in the domain of face recognition.

"Adapt GWO for Face Recognition problem and evaluate its performance with other meta-heuristic algorithms."

1.4 Scope of work

GWO is inspired from the social behavior of grey wolves. They are divided into four types of groups according to their behavior to perform food searching. The four groups are termed as Alpha, Beta, Gama and Omega. They represent the optimality of solution in GWO; Alpha represents the best and Omega represents the least. GWO uses optimization capability to find a best match of input image in the database.

We applied the GWO in face recognition and compare it with other nature inspired algorithms like ACO, BBO, PSO, and Extended BBO based on accuracy of recognition and time taken to recognize input image. We used MATLAB programming platform for this purpose. Here ORL standard database by Olivetti Research Laboratory in Cambridge, U.K and Cohn-kanade AU- coded facial expression by Kanade, T.Cohn, J.F., & Tian, Y. is used.

Hence scope of our work can be summarized as:-

1. Adapt GWO for face recognition problem.
2. Empirical study of GWO for face recognition problem on the basis of accuracy and time.
3. Comparing the results of GWO, BBO, ACO, PSO and extended BBO.

1.5 Organization of thesis

Rest of our work can be summarized as below:-

Chapter 2 reveals about the previous research work done in the field of face recognition

Chapter 3 contains the details of metaheuristic algorithms like ACO, BBO, PSO and extended BBO in solving face recognition problem.

Chapter 4 explains the usage of Grey Wolf Optimization in solving face recognition problem.

Chapter 5 illustrates the working of our experiments and comparison with previous approaches.

Chapter 6 concludes our work and explained about its future work.

CHAPTER 2

LITERATURE REVIEW

Face Recognition becomes one of the biometrics authentication techniques used from the past few years. Face recognition consists of Pattern recognition and Image analysis. Face recognition encompasses two main tasks: verification and identification. Former is a 1:1 match i.e. an images against a template face images while later is a 1:N problem that compares a input face image against all image templates of a database. Facial perception is daily task of human which shows their capability then why can't be a such computer system for human face perception capability be possibility for authentication problem in our neighborhood. Automatic face recognition really started by Kelly [23] and Kanade [29].

Face recognition is used in various fields like security, authentication, gaming etc. some of the real world example are shown in the following table.

Commercial products	Websites
FaceIt from Visionics	http://www.FaceIt.com
Viisage Technology	http://www.viisage.com
FaceVACS from Plettac	http://www.plettac-electronics.com
FaceKey Corp	http://www.facekey.com
Cognitec Systems	http://www.cognitec-systems.de
BioID sensor fusion	http://www.bioid.com
Biometric Systems, Inc.	http://www.biometrica.com

Table 1 Available Commercial Face Recognition Systems

A typical face recognition process is illustrated in block diagram:-

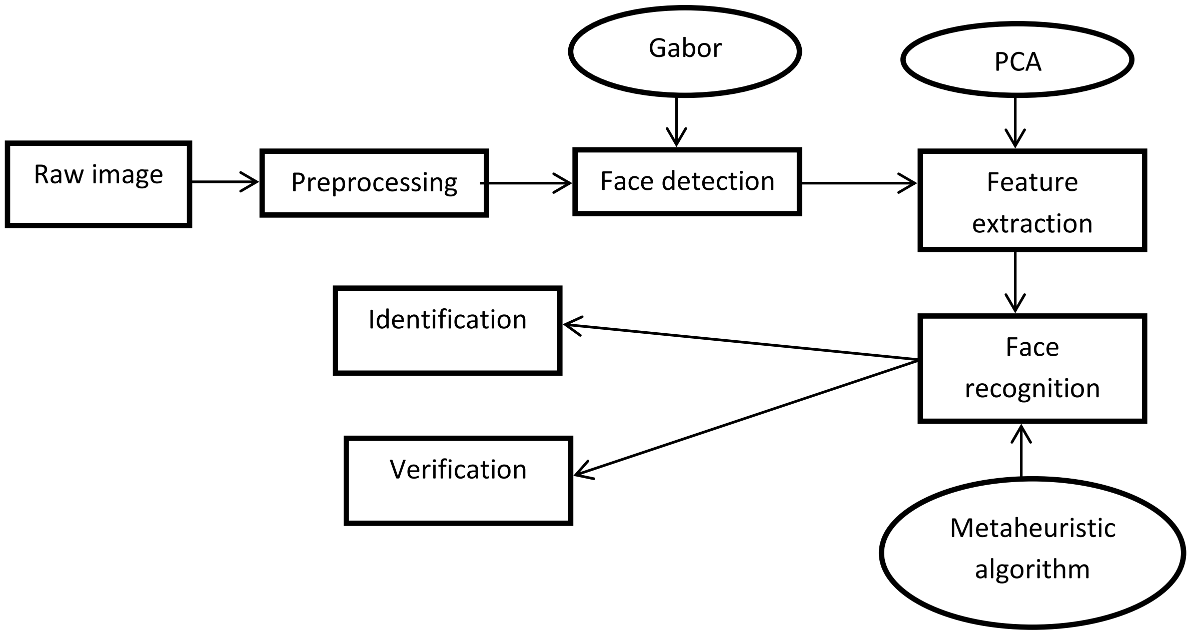


Figure 1 generic face recognition proces

2.1 Steps of face recognition

The process of face recognition can be divided three steps:-

(1) Detection and normalization of faces,

(2) Feature extraction from faces,

(3) Recognition by identification and matching.

2.1.1 Face Detection

The primary goal of this step is to find whether a human face exists in the image or not [9]. If a human face exists in the image given then it gives the location of face in that image. If more than one faces are found than it gives the templates defining each face as expected output. Alignment of these templates is performed in order to make face recognition system more secure. Face detection may also be used in classification of video streams. . The main goal of face detection is to find faces in the image and returns the location of face in the image. Pre-processing may be required to achieve better result like remove the noise. Pose presence, Facial expression, Image orientation are the various terms that affects the face detection. The main techniques used for this purpose are featured based approach and appearance based approach [16].

2.1.2 Feature extraction and selection

After detecting the face in image, features are collected that are fed into a face recognition system. According to the type of recognition system, features can be local such as lines or fiducial points, or facial such as eyes, nose, and mouth. Features may also be used in face detection; in that case features are extracted simultaneously with face detection. Feature extraction is also used in animation and recognition of facial expressions.

These templates cannot be used directly as they may contain thousands of pixels and are not better to make a recognition system. Moreover the templates may differ in some aspects like illumination, alignment and occlusion and clutter may also be there. Sometimes, several steps like Face detection and feature extraction can be processed at a time , e.g. facial features like eyes, nose, mouth used in both steps.

The important features that are enough to represent a face are selected in this step. In our approach we use Principle Component Analysis for extracting the features from input image. Optimal features from the extracted features to increase the efficiency and decrease the overhead of input feature. We have used Particle Swarn Optimizaton for feature selection in our approach.

2.1.3 Face recognition

There is a database where set of images are stored and presence of input image is to be recognized. When optimal features are selected from the list of extracted features, Face recognition is performed to match the input image with the database. To automate the recognition system database is built in which images are stored after extracting and selecting the features. These features are matched with the input image features. Based on this match report, classification is done to identify the portion to which image belongs.

Identification and verification are the two main application of face recognition. Identification gives the output as a similar image as input image while verification gives output as YES or NO depending on the presence or absence of input image in the database. Following figure depicts these steps.

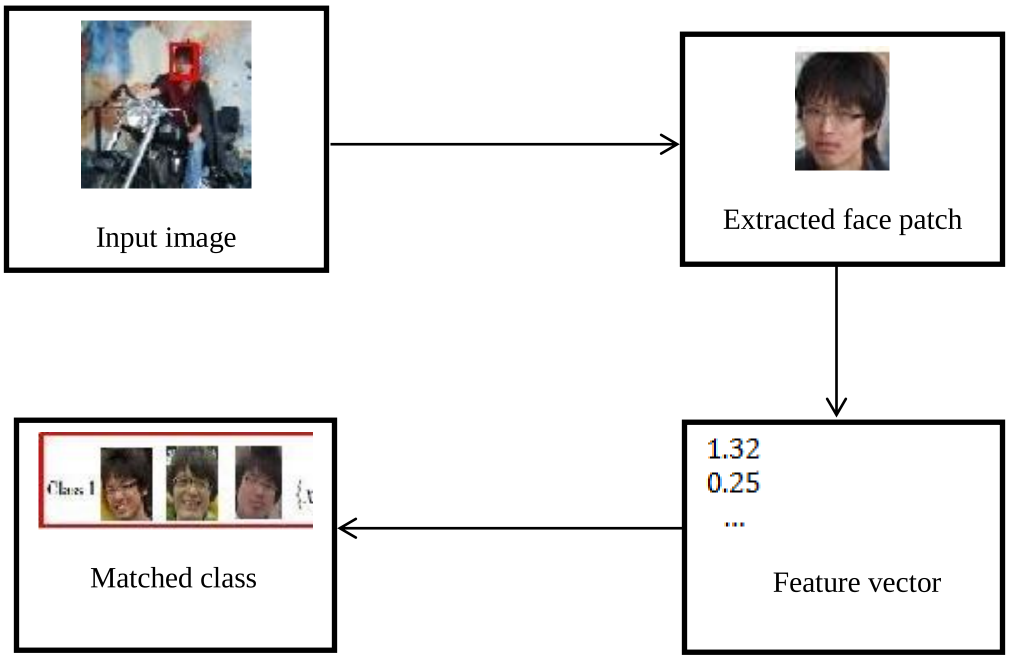


Figure 2 example of recognition process

There are three approaches for face recognition [7]:-

A. Feature based approach

Local features of faces like nose, eyes are segmented and used as input in face detection to perform task of face recognition.

B. Appearance based approach

Appearance based approach approach uses whole face as a input to the face detection system to process face recognition.

C. Hybrid approach

This is the combination of both feature based and holistic approach. Both local features of face and whole face is used as input to face detection system.

We are interested in only two phases feature extraction and selection , and face recognition. Some of the techniques used in these phases are explained below-

2.2 Principle Component Analysis

Principal component analysis (PCA) is a statistical procedure. PCA is mathematically defined [7] as an orthogonal linear transformation that transforms the data to a new coordinate system such that the greatest variance by some projection of the data comes to lie on the first coordinate (called the first principal component), the second greatest variance on the second coordinate, and so on.

Basically reduction in dimension of feature space is done by the Principle Component Analysis that represents our data well. This reduction minimizes the computational cost [20]. Following figure shows how the principle components are selected. In figure X-Y coordinate is used to represent the dataset initially. We found principle dimensions as U-V by using PCA as new dimension have the minimum correlation. In this case mapping is done from X-Y axis to U-V axis and the resulted covariance becomes zero. We can say the U-V axis is the principle directions (or they may be components in other cases).

2.2.1 PCA in face recognition

Turk [20] proposed the face recognition method was proposed based on Principal Component Analysis (PCA). From a set of the training faces a subset of principal directions (principal components) selected by this method. Then projection of these components in the space gives the feature vectors. Different distance measures are used to compare these feature vectors in order to perform face recognition. In the PCA-based face approach, eigenvectors and eigenvalues of the covariance matrix of the training data are calculated. Calculation of eigenvectors may be complicated in case of very large matrix.

PCA basically used in feature extraction phase of face recognition. An image can be represented as a two dimensional array of pixel value in space. PCA helps us to find the optimal subset of these pixels used to represent the image well. Following steps represents the PCA [6] method.

1) Take two dimensional array (N * N) corresponds to sample images and convert into gray scale. Now 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 of $N^2 *1$. Store the variation of each face to mean in matrix B.

3) Calculate covariance matrix $C=BB^T$ of N^2*N^2. For dimension calculate covariance matrix 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 $u_1,u_2,u_3\ldots\ldots\ldots\ldots..u_m$.

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 Gabor Filter

A good and proper image is required to perform efficient face recognition, so there may be a need of preprocessing. Contrast and illumination is aligned after the grey scale conversion of the images. Gabor filter is used to select the optimal features from the images. Gabor filter is a band-pass linear filter which is defined by multiplication of a harmonic function multiplied by Gaussian function. Thus, a complex sinusoidal plane of particular frequency and orientation modulated by a Gaussian envelope represented gabor filter. Moreover 2D kernel can be expressed as below [10,17] :-

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

Where

$t = a,,\cos \theta_v + b,,\sin \theta_v$

$d = -a,,\sin \theta_v + b,,\cos \theta_v$

$f_u = f_{max}/2(u/2)$ and $\theta_v = v\pi/8$

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

f_u and θ_v represents Central frequency and orientation respectively. And ratio between central frequency and the size of Gaussian envelope calculated by h and p. Different values of parameters mentioned above can be used for evaluating characteristics of the kernels, the most common parameters for the face recognition are $\kappa = \eta = \sqrt{2}$ and $f_{max} = 0.25$ [20,12].

2.4 Classifier

Most commonly used classifier is Minimum Distance Classifier which follows Euclidean distance for identifying the similarity between two patterns and is also a popular method to solve pattern matching problems. Let x be an unknown pattern to be classified and zi (i = 1, ..., n) be a prototype for category ωi. x and z are m-dimensional vectors in the feature space, n is the number of categories, and m is the number of dimensions of the feature space. The MDC is defined as:

$$x \in \omega i,\ if\ \ dist(x, zi) = min\{dist(x, zj)\},\ \forall j$$

where dist(·) is the Euclidean distance function

$$(x, zi) =[\Sigma^m{}_{k=1}(x_l - z_{ik})^2.]^{1/2}$$

CHAPTER 3

GENERIC EVOLUTIONARY TECHNIQUES

This chapter explain the various evolutionary algorithms used in optimization problems so far. There are numerous algorithms that are inspired by social behavior of species found in the world and uses their social behavior in order to search optimal solution of optimization problems. Ant Colony Optimization(ACO), Particle Swarm Optimization(PSO), Biogeographical Based Optimization(BBO), Extended Species Abundance Model of Biogeography(Extended BBO) and Grey Wolf Optimization(GWO) are some of the examples of generic evolutionary algorithms inspired by social behavior of species and are explained below.

3.1 Ant Colony Optimization

The first ACO algorithm is introduced by Marco Dorigo and colleagues in the early 1990's. ACO works on the basis of social behavior of Ants in their colonies to search for food. Basically an Ant has limited capability but when we talk about their colony their searching behavior is more interesting. Pheromone is spread out on the ground by the Ants in order to mark some optimal path that if followed by other members may be good for food searching. Ant colony optimization works on a similar approach for solving optimization problems.

In ACO, a number of artificial ants build solutions to an optimization problem and exchange information on their quality via a communication scheme that is reminiscent of the one adopted by real ants [12].

At the beginning pheromone is of very low level is spread out by ants. Each ant applies a state transition rule, which is-

$$S=\begin{cases} \arg max_{u\epsilon j(r)} \{[t(r,u)].[\eta(r,u)^{\beta}]\} & \text{if } q \leq q0 \ \text{(exploitation)} \\ S, & otherwise\ (biased\ exploitation) \end{cases}$$

Where:-

(r, u) = an edge between point r and u,

t(r, u) = pheromone on edge (r, u).

h(r, u) = desirability of edge (r, u), which is usually defined as the inverse of the length of edge

q = a random number uniformly distributed in [0, 1],

q0 = user-defined parameter with ($0 \le q_0 \le 1$),

b = parameter controlling the relative importance of the desirability.

Pheromones are updated upon each walk of ant on a way y some rules which are as follows.

Local updating rule: While constructing its tour, every ant will modify the pheromone value on the travelled edges this rule.

$$\tau(r,s) \leftarrow (1-\rho).\tau(r,s) + \rho.\tau_0$$

where ρ represents pheromone evaporation (note: $0 \le \rho \le 1$).

Global updating rule: after reaching at the destination, pheromone on the edge is modified again by this rule.

$$\tau(r,s) \leftarrow (1-\alpha).\tau(r,s) + \alpha.\Delta\tau(r,s)$$

3.1.1 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 an optimal 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 [8, 30] . The problem is represented in form a digraph with only 2n or O(n) arcs instead of a complete graph with $O(n^2)$ 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 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.

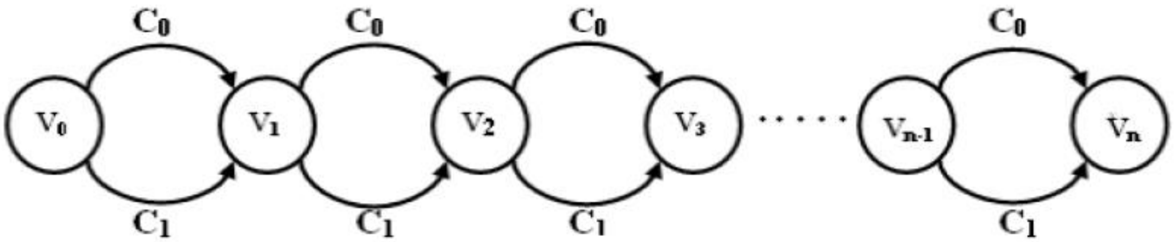


Figure 3 The diagraph

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 $f_1, f_2, \ldots. f_n$. In graph the i^{th} node v_i is used to represent feature f_i. 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 chosen 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.

$$p_i^j(t) = \frac{\left[\tau_i^j(t)\right]^{\alpha}\left(\eta_i^j\right)^{\beta}}{[\tau_i^0(t)]^{\alpha}\left(\eta_i^0\right)^{\beta} + [\tau_i^1(t)]^{\alpha}(\eta_i^1)^{\beta}} \qquad (i = 1,2 \ldots\ldots.n; \qquad j = 0,1)$$

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

$$F(s) = \frac{N_{corr}}{1+\lambda N_{feat}}$$

where N_{corr} the number of examples that are correctly classified, N_{feat} 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 pheromones on path traversed by best ant (Local updation). Update pheromones of every path (evaporation). When all the ants have completed a solution, the trails are updated by

$$\tau_i^j(t+1) = \rho.\tau_i^j(t) + \Delta\tau_i^j(t) + Q_i^j(t)$$

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.

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

3.2 Particle Swarm Optimization

Eberhart and Kennedy, in 1995 proposed Particle Swarm Optimization (PSO) based on social behaviors of flock of birds and is biologically inspired computational search and optimization method. PSO works as a parallel search technique in which particles are the multi-dimensional search space. At any particular instant, each particle has a position and a velocity. The position vector of a particle against the origin of space denotes the trial solution of the searching problem. At the starting point of optimization, a particle population is initialized with random velocities v_i and random location x_i. The population of such particles is called a “swarm” S. PSO works exactly as birds behaves in search for their food in their swarm.

Bird while in their swarm tries to choose the optimum path towards the food by choosing best of neighbor's path to the food. This is done by the location and velocity towards the food. The velocity also depends on the previous local best and global best of the swarm.

Each particle works in collaboration with its neighbors in order to find the optimal path o the food. Basic algorithm by Kennedy and Eberhart (1995) given as below:-

X^i_k - particle position

V^i_k - particle velocity

P^i_k - previous best individual paricle position

P^g_k - previous best swarm postion

C_1, C_2 - cognitive and social parameters

R_1, R_2 - random numbers varying from 0 to 1.

Individual particle position will updated as:

$X^i_{k+1} = X^i_k + V^i_k+1,$

And the velocity as:

$V^i_{k+1} = V^i_k + C_1R_1(P^i_k - X^i_k) + C_2R_2(P^g_k - X^i_k).$

General algorithm for PSO can be illustrated as below:-

- Initialize particles with their random location and velocity.
- Fitness function is used to calculate the fitness value of each particle. If it is better than previous then update it as local best.
- When all particle's local best are updated, update the global best of swarm.
- Update the velocity and location of each particle using local and global best.
- If maximum iteration are not completed repeat from step 2.

3.2.1 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 [26] 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 adaption of 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 represents 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.

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.

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 neighboring 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. [11]. The variables that define habitability are known as Suitability index variables (SIVs). SIVs are the independent variables but the HSI of each habitat depends on SIVs. Therefore habitats with high HSI value have 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 tend 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 comparison to the habitats with low HSI.

We plot a graph as shown in Figure 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 number of 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 rise. When number of species are between S_1 and S_2

there exist 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.

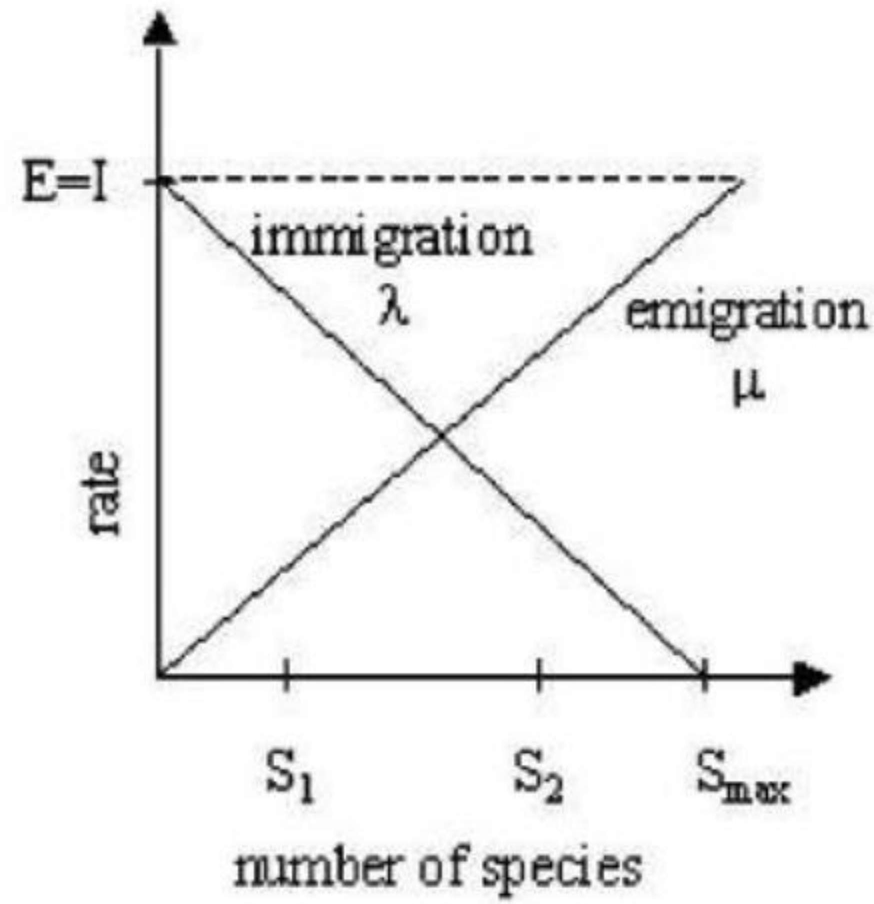


Figure 4 species abundance in Single Island

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 chosen 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 chosen 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 counts 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.

3.3.1 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 an optimal feature vector set [5].

In the recognition phase we have used biogeography based optimization to recognize an input test image. 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.

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 an 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.

3.4 Extended Species Abundance Model Of Biogeography (Extended BBO)

Extended BBO was proposed by [4], 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 nonzero 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[4].

Decline rate = Emigration rate + Extinction rate = $Ek^{\Psi}/_{n} + \sqrt{nk}$; where Ψ is dependency factor, E is emigration rate.

Growth rate= Immigration rate + Evolution rate = I (1 - $\frac{k}{n}$) + $\frac{c}{k}$; where c is constant, I immigration rate, k is rank of SIV, n is total number of valid SIV.

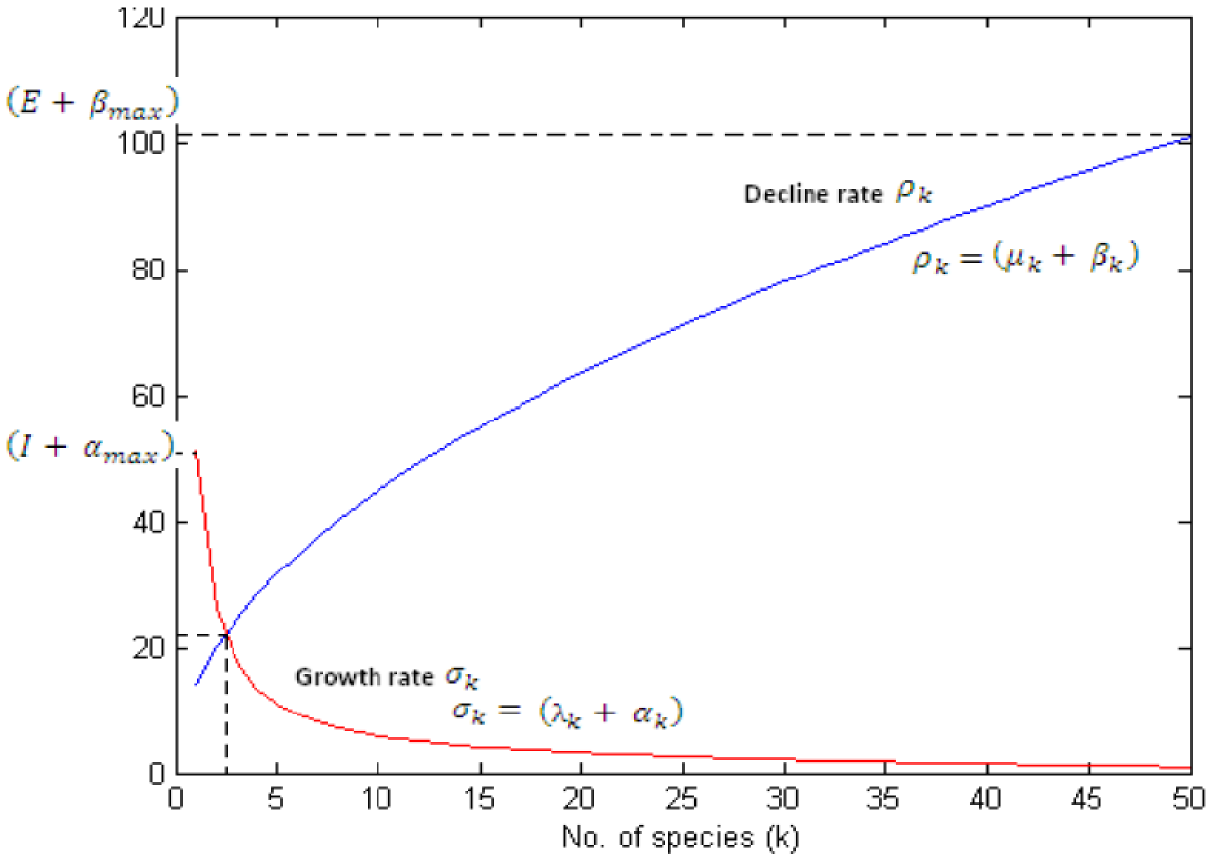


Figure 5 Extended Species Abundance Model

As we can see extended species abundance model as well as probability curve is same in all linear and non-linear migration models [27]. 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.

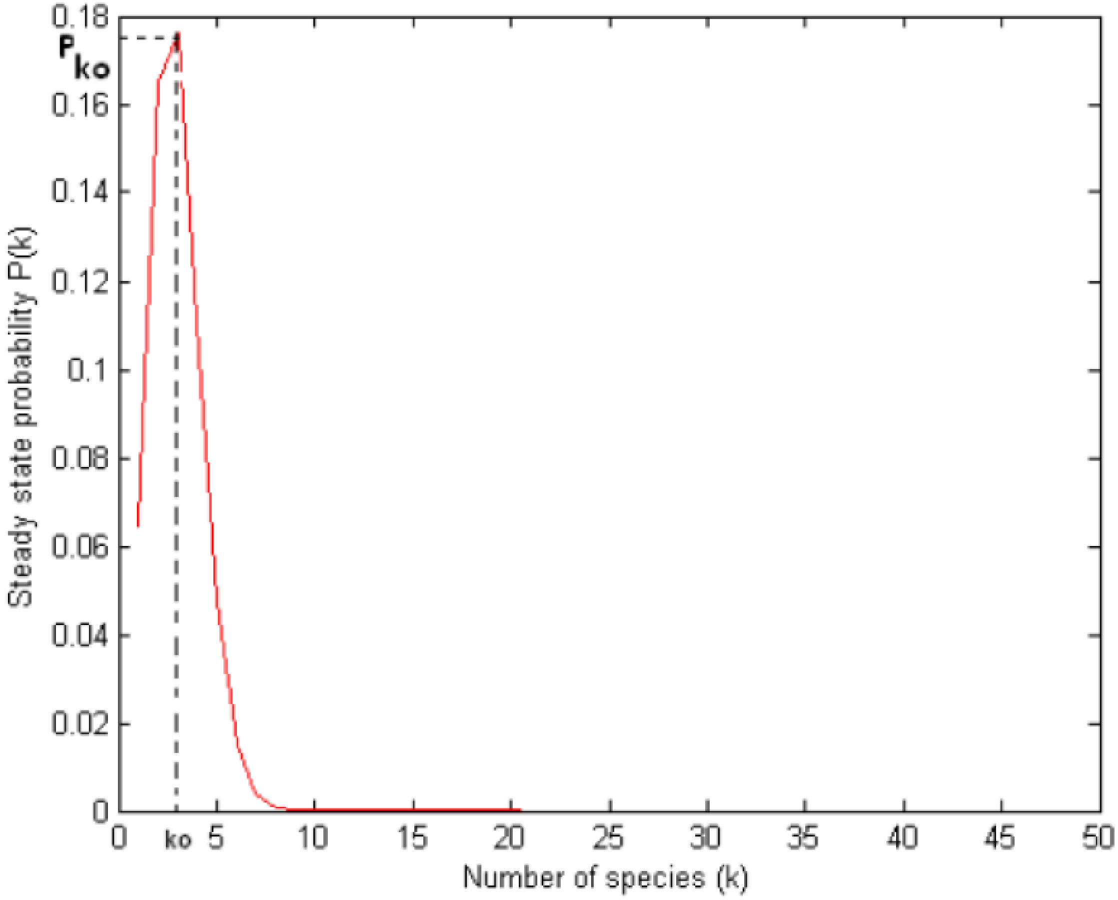


Figure 6 Steady state probability curve P_k as function of species k

In abundance model presented above, maximum value of immigration and emigration are I and E respectively. K0 is equilibrium point where emigration and immigration rates are equal.

3.4.1 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 an 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.

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.

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. These images constitute 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.

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.

3.5 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 an optimal feature vector set. In the recognition phase we used particle swarm optimization to reduce the time consumed 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.

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.

CHAPTER 4

GWO DESCRIPTION

This work proposes another rising meta-heuristic called Grey Wolf Optimizer (GWO) motivated by grey wolves (Canis lupus) in the field of Face Recognition. The Grey wolf Optimizer algorithm impersonates the administration progression and chasing system of grey wolves in nature. There are four sorts of grey wolves, for example, alpha, beta, delta, and omega are utilized for reenacting the initiative chain of command. Furthermore, the three fundamental strides of chasing, hunting down prey, enclosing prey, and assaulting prey, are actualized. The technique is then benchmarked on twenty nine test parameters, and the outcomes are checked by a relative study with, Gravitational Search Algorithm (GSA), Evolutionary Programming (EP), Particle Swarm Optimization (PSO), Differential Evolution (DE) and Evolution Strategy (ES) [2].

4.1 Inspiration

Fundamentally Grey wolf (Canis lupus) fits in with Canidae group. They are considered as pinnacle predator, which implies that they are at the highest point of the natural food chain. Generally grey wolves like to live in a pack. The gathering size is 5–12 overall. Exceptionally compelling is that they have an extremely strict social prevailing progression as demonstrated in Figure.

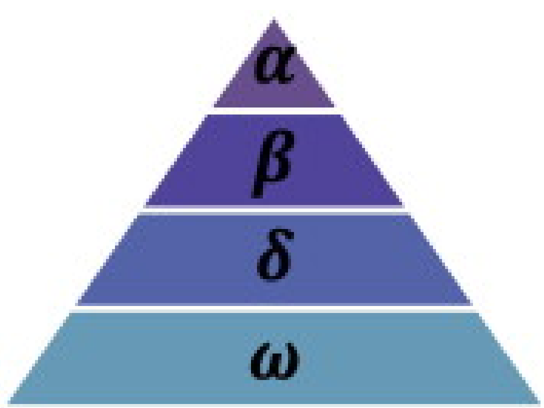


Figure 7 Hierarchy of Grey Wolf

In the gathering of Grey wolves, the pioneers can be male and a female, called Alphas. Alphas are leaders. The alpha is for the most part in charge of settling on choices about chasing, resting spot, time to wake, etc. The alpha's choices are managed to the pack. In any case, a just conduct has likewise been seen, in which an alpha takes after alternate two-timers. In the group, the whole pack recognizes the alpha by holding their tails down. The alpha ones are additionally called the prevailing wolf since his/her requests ought to be trailed by the pack. The alpha wolves are just permitted to mate in the group. Interestingly, the alpha is not so much the most grounded individual from the pack but rather the best regarding dealing with the pack. This demonstrates that the association and control of a pack is considerably more vital than its quality. Accordingly, Alpha wolves are the prevailing two-timer. Thus they are at the top in the chain of command of dark wolves.

The second level in the chain of command of wolves is Beta. The betas are subordinate ones that help the alpha in choice making or other pack exercises. The beta wolf can be either female or male, and he or she is most likely the best possibility to become the alpha. They obtain the position of alphas in the event that one of the alphas passes away or turns out to be extremely old. The beta wolf ought to regard the alpha, however orders the other lower-level wolves too. It assumes the part of a consultant to the alpha and the discipliner for the pack. The beta fortifies the alpha's orders all through the pack and offers input to the alpha.

The most reduced in the progression of grey wolf is omega. The omega assumes the part of substitute. Omega wolves dependably need to submit to the various overwhelming wolves. They are the last ones that are permitted to eat. It may appear the omega is not a vital individual in the pack, but rather it has been watched that the entire pack face inside battling and issues if there should arise an occurrence of losing the omega. This is because of the venting of viciousness and dissatisfaction of all wolves by the omega. This helps fulfilling the whole pack and keeping up the predominance structure. At times the omega is additionally the sitters in the pack [2].

The third position in the pecking order goes to Delta. On the off chance that a wolf is not an alpha, beta, omega, then he/she is called as subordinate. Delta wolves need to submit to alphas and betas, however they overwhelm the omega. Scouts, sentinels, seniors, seekers, and overseers have a place with this classification. Scouts are in charge of viewing the limits of the domain and cautioning the pack in the event of any peril. Sentinels ensure and guarantee the well-being of the pack. Seniors are the accomplished wolves who used to be either alpha or beta. Seeker's help the alphas and betas when chasing prey and giving food to the pack. At long last, the guardians are in charge of watching over the frail, sick, and injured wolves.

In the social order of wolves, group chasing is another intriguing social conduct of grey wolves. As per Muro et al., the fundamental periods of grey wolf chasing are as follows [2]:

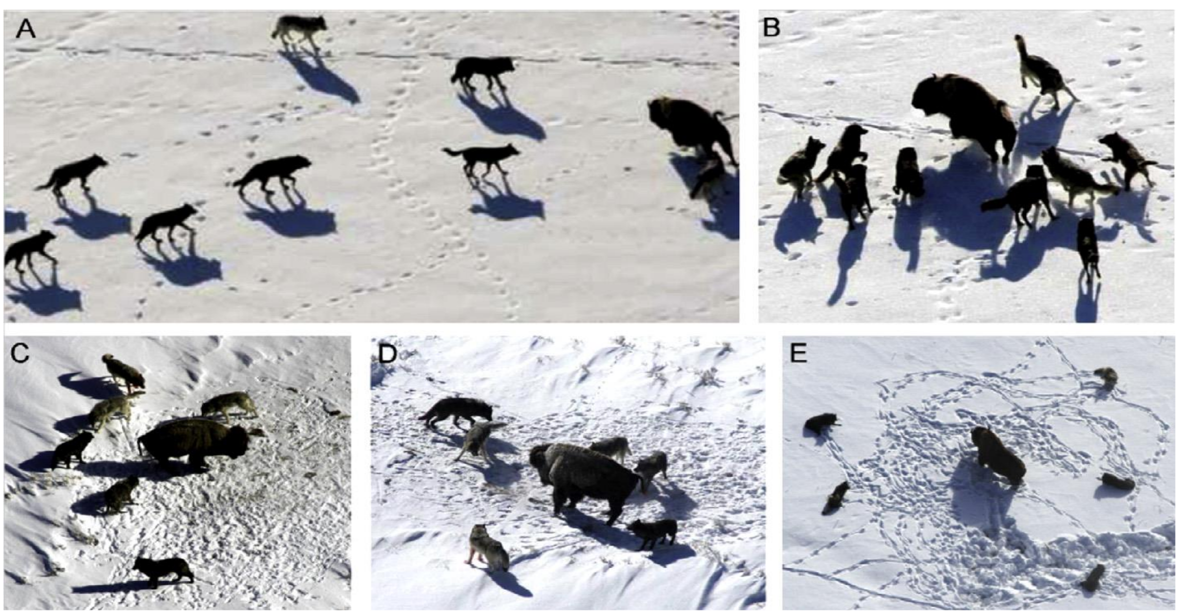


Figure 8 Hunting Behaviour of Grey Wolves [2].

- Tracking, pursuing, and drawing closer the prey.
- Pursuing, circling, and pestering the prey until it quits moving.
- Attack towards prey.

These above steps are shown in Fig 3. In this work this chasing method and the social pecking order of grey wolves are numerically displayed keeping in mind the end goal to outline

GWO and perform enhancement.

4.2 Mathematical model and algorithm

In this section the mathematical models of the social hierarchy, tracking, circling, and assaulting prey are provided.

4.2.1 Social hierarchy

In the social progressive system of wolves when planning GWO, we consider the fittest arrangement as the alpha (α). Thus, the second and third best arrangements are named beta (β) and delta (δ) separately. Whatever is left of the applicant arrangements are thought to be omega (ω). In the GWO technique, the chasing (hunting) is guided by α, β, and δ. The ω wolves take after these three wolves [2].

4.2.2 Encircling prey

As mentioned above, during the chase, wolves encircle prey. In the mathematically model encircling conduct the following equations are proposed:

$$\mathbf{D = | C. _p \, \boldsymbol{X_p}(t) - A. X (t)|}$$

$$\mathbf{\boldsymbol{X} (t+1) = X_p(t) - A. D}$$

where t demonstrates the current iteration, A and C are coefficient vectors,
ω_p is the position vector of the prey, and ω demonstrates the position vector of a grey wolf.

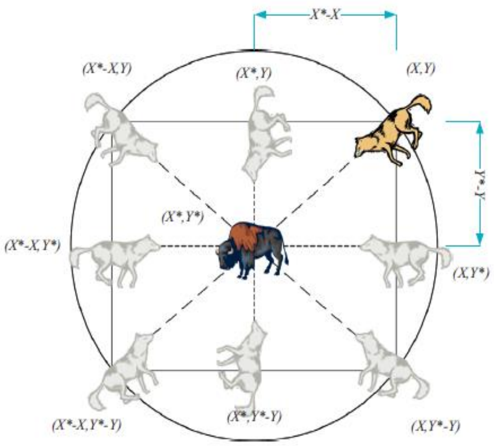
The vectors A and C are evaluated as follows:

$$\mathbf{A = 2\alpha \, . \, r_1 - \alpha}$$

$$\mathbf{C = 2 \, . \, r_2}$$

where components of α are straightly diminished from 2 to 0 through the span of iterations and r_1, r_2 are arbitrary vectors in [0, 1].

To see the impacts of initial two mathematical statements, a two-dimensional position vector and a portion of the conceivable neighbors are given in Fig. 4(A). As can be found in this figure, a grey wolf can overhaul its position(X, Y) as indicated by the position of the prey (X*, Y*).



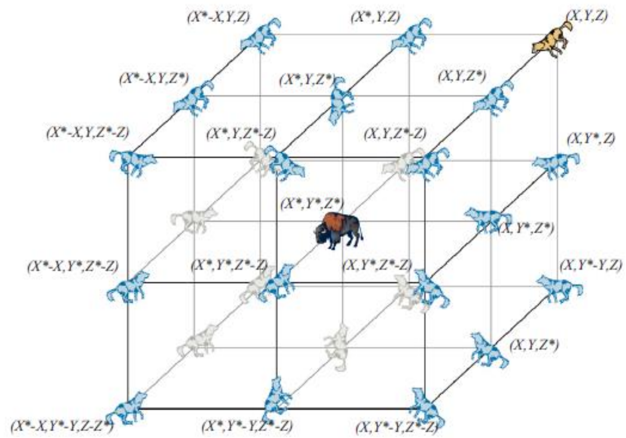


Figure 9 2D & 3D position vectors and their possible next locations.

Better places around the best agents can be come to regarding the present position by conforming the estimation of A and C vectors. For example, (X*–X, Y*) can be come to by setting A=(1,0) and C=(1,1). The possible upgraded positions of a grey wolf are illustrated in Fig. 4(B). Note that the arbitrary vectors r_1 and r_2 permit wolves to achieve any position between the points showed in Fig. 3. So a grey wolf can overhaul its location inside the area around the prey in any irregular area by using first two equations.

The same idea can be stretched out to an inquiry space with m dimensions and the wolves will proceed in hyper-solid (or hyper-cubes) around the best solution got as such.

4.2.3 Hunting

Grey wolves can perceive the area of prey and enclose them. The chase is normally guided by the alpha wolves. The beta and delta may likewise take part in chasing occasionally. On the other hand, in an abstract search space we don’t have clue about the area of the ideal (prey). With a specific end goal to mathematically simulate the chasing conduct of grey wolves, we assume that the alpha (best solution), beta, and delta have better learning about the potential area of prey. Hence, we spare the initial three best solutions got so far and oblige the other agents (counting the omegas) to redesign their positions as per the position of the best hunt agents. The following formulas are proposed in this context.

$$\mathbf{D_a = |\, C_1 \,.\, X_a - X \,|,}$$
$$\mathbf{D_b = |\, C_2 \,.\, X_b - X \,|,}$$
$$\mathbf{D_c = |\, C_3 \,.\, X_c - X \,|,}$$

$$\mathbf{X_d = X_a - A_1 \,.\, (D_a),}$$
$$\mathbf{X_e = X_b - A_2 \,.\, (D_b),}$$
$$\mathbf{X_f = X_c - A_3 \,.\, (D_c),}$$

$$\mathbf{X\,(t+1) = (\,X_d + X_e + X_f) \,/\, 3,}$$

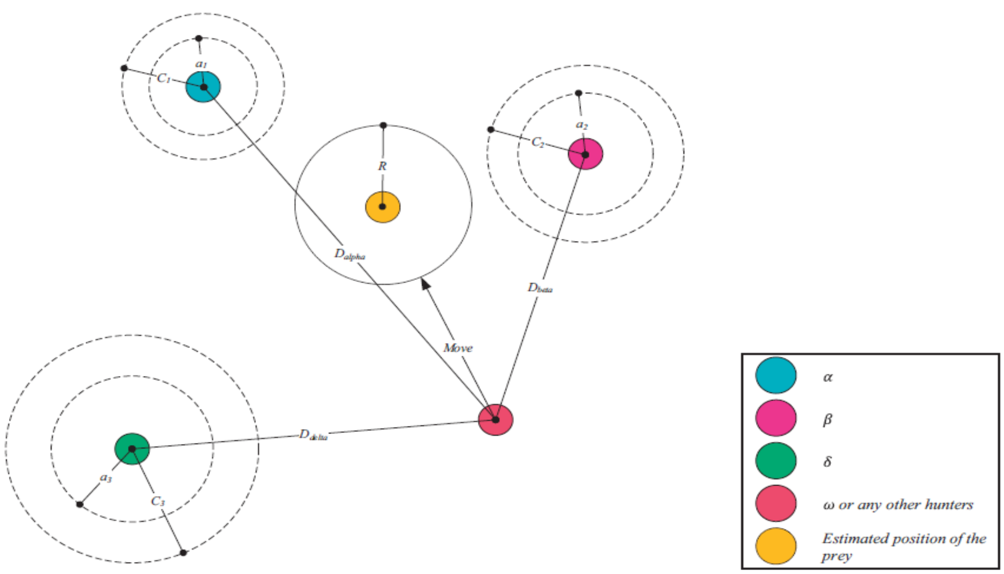


Figure 10 Position updating in GWO.

Fig. 10 shows how an search agents upgrades its position as per alpha, beta, and delta in a two dimensional search space. It can be watched that the last position would be in an arbitrary place inside of a circle which is characterized by the location of alpha, beta, and delta wolves in the pursuit space. However, alpha, beta, and delta appraise the position of the prey, and other grey wolves overhaul their positions arbitrarily around the prey

4.2.4 Attacking prey (exploitation)

As said over the grey wolves complete the chase by assaulting the prey when it quits moving. According to mathematical model, drawing nearer the prey we diminish the estimation of a. Note that the fluctuation range of A is additionally diminished by a. However, A will be an arbitrary value in the interval [-2a, 2a] where a is diminished from 2 to 0 through the span of iterations. When random values of A are in [-1, 1], the next location of a search agent can be in any position between its present position & position of the prey.

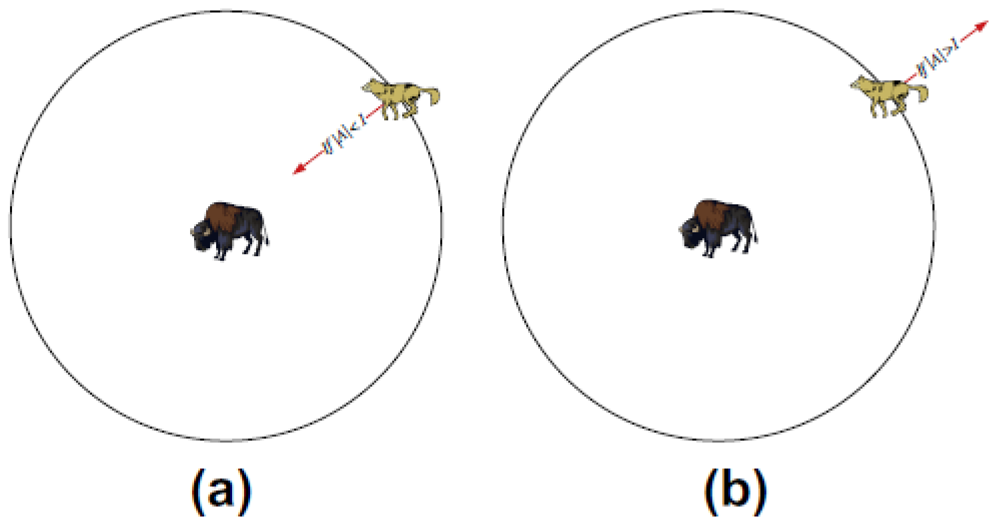


Figure 11 Attacking prey versus searching for prey.

Fig. 11(a) demonstrates that |A| < 1 powers the wolves to assault towards the prey. With the operators proposed as such, the GWO algorithm permits its search agents to redesign their position taking into account the area of the alpha, beta, and delta; and assault towards the prey. Thus, the GWO algorithm is inclined to stagnation in neighborhood solutions with these operators. Circling mechanism shows investigation to some degree, yet GWO needs more emphasize to underscore investigation [2].

4.2.4 Search for prey (exploration)

Grey wolves search as per the current position of the alpha, beta, and delta wolves. They separate from one another to hunt down prey and join to assault prey. In the mathematical model, we use A with arbitrary values more than 1 or not exactly -1 to oblige the search agent to veer from the prey.

This stresses investigation and permits the GWO algorithm to pursuit all inclusive. Fig. 11(b) likewise demonstrates that $|A| > 1$ strengths the grey wolves to veer from the prey to ideally locate a fitter prey. Another segment of GWO that supports investigation is C. The C vector contains irregular values in [0, 2]. This segment gives irregular weights to prey keeping in mind the end goal to stochastically underline ($C > 1$) or deemphasize ($C < 1$) the impact of prey in characterizing the separation in. This helps GWO to demonstrate a more arbitrary conduct all through optimization, favoring investigation and neighborhood optima avoidance. It merits saying here that C is not directly diminished as opposed to A. We intentionally oblige C to give arbitrary qualities at all times to underscore investigation not only during initial one but also final iteration.

This part is extremely useful if there should arise an occurrence of nearby optima stagnation, particularly in the last iterations. The C vector can be additionally considered as the impact of obstacles to drawing nearer prey in nature. As a rule, the obstacles in nature show up in the chasing ways of wolves and drawing closer prey. This is precisely what the vector C does. According to the position of a wolf, it can arbitrarily give the prey a weight and make it harder and more remote to reach for wolves, or vice versa [2].

To entirety up, the search procedure begins with making an arbitrary populace of grey wolves in the GWO algorithm. Through the span of iterations, alpha, beta, and delta wolves gauge the likely position of the prey. Every competitor solutions redesign its separation from the prey. The parameter a is diminished from 2 to 0 to underline investigation and abuse, respectively. Candidate solutions have a tendency to veer from the prey when $|A| > 1$ and merge towards the prey when $|A| < 1$. At last, the GWO algorithm is ended by the fulfillment of an end standard.

4.3 GWO in face recognition

The pseudo code of grey wolf optimizer is mentioned below:

- The proposed social progressive system helps GWO to spare the best solutions got so far throughout cycle.
- The proposed circling system characterizes a circle-shaped neighborhood around the solutions which can be reached out to higher dimensions as a hyper-circle.
- The random parameters 'A' and 'C' help competitor solution to have hyper-circles with diverse arbitrary radii.
- The proposed chasing technique permits competitor solutions to find the probable position of the prey.
- Exploration and abuse are ensured by the versatile estimations of 'a' and 'A'.
- The versatile estimations of parameters 'a' and 'A' permit GWO to easily move in the middle of exploration and abuse.
- With diminishing 'A', a half of the cycles are dedicated to exploration ($|A|>1$) and the other half are committed to exploitation ($|A| < 1$).
- The GWO has just two fundamental parameters to be adjusted (a, C).

More clearly we can illustrate the process of Grey Wolf in Face Recognition as below:-

Input: n Number of grey wolves in the pack,

Stopping Criteria: Number of iterations for optimization

Output: x_α which represents optimal grey wolf position.

I. Initialize a population of n grey wolves' positions randomly.

II. Find the α, β and δ solutions based on their fitness values.

III. While Stopping criteria not met do

a. For each Wolf in pack

b. do

Update current wolf's position according to $X(t+1)=(X_\alpha + X_\beta + X_\delta)/3$.

c. end

d. Update a, A, and C:

e. Evaluate the positions of individual wolves

f. Update α; β; and δ:

IV. End While

4.3.1 Description of algorithm

1. To the training set of images apply PCA.
2. PCA generates a set of eigenfaces, say k lesser than the size of the original training set, which are also face images but represent just the major features of a training set.
3. Each face image is represented as some proportion (weight) of each eigenface forming a weight vector,

$$F_i = [w_1, w_2, w_3 \dots w_k]$$

4. On the eigenfaces generated by PCA apply the GWO technique as follows:

    - Initialize the population randomly for possible solution.
    - Calculate the fitness of each search agent.
    - Calculate the fitness function of each search agent X_α, X_β, X_δ
    - For each search agent, update the position of current search agent by

$$X(t+1)=(X_\alpha + X_\beta + X_\delta)/3.$$

    - Update a, A and C; calculate fitness of all search agent

where F_i is the i^{th} face vector.

- Update X_α, X_β, X_δ
- Return the image with optimal value X_α.

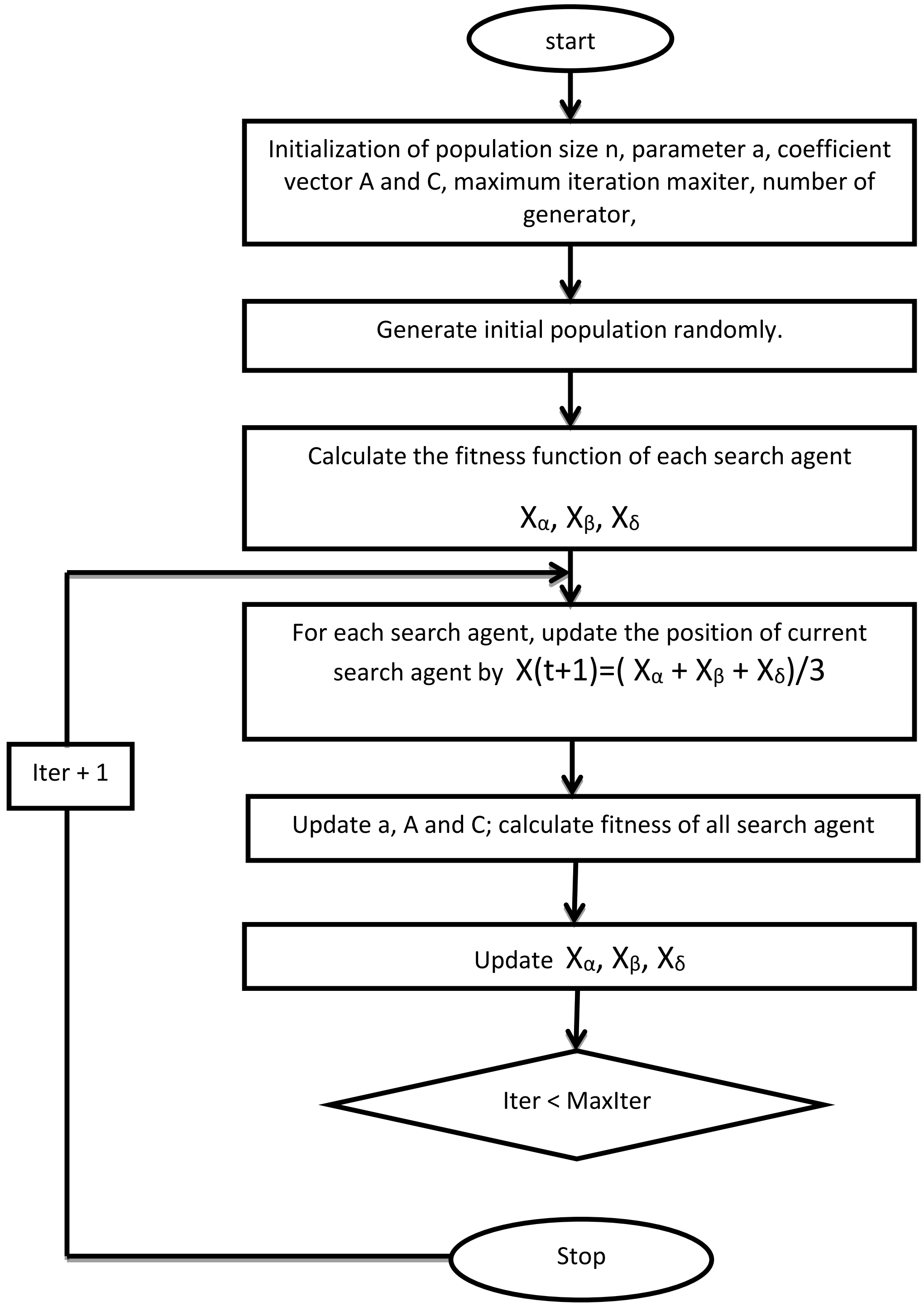


Figure 13 - flow chart for Grey Wolf Optimization in Face Recognition

Several assumptions regarding the terminology related to Grey Wolf Optimizer in face recognition process include:-

Wolf:- If we assume that a wolf represents a single solution, we can give wolf the following properties.

- Each wolf is separate independent potential solution represented by one individual in the population.
- In our case wolf represents feature vector of a single image.
- Difference between the positions of wolves decides the size of steps.
- Every image can be match with other images according to difference between the feature vector of images.

Pack:- In GWO, the following features can be imposed concerning a pack:

- The numbers of wolves are fixed.
- In our case pack represents a complete set of image matching.
- A pack is an individual of the population and the number of wolves in the pack is equal to the size of the population
- Alpha wolf is represented by the fittest solution in the iteration.

CHAPTER 5

MODEL OF FACE RECOGNITION

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 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.

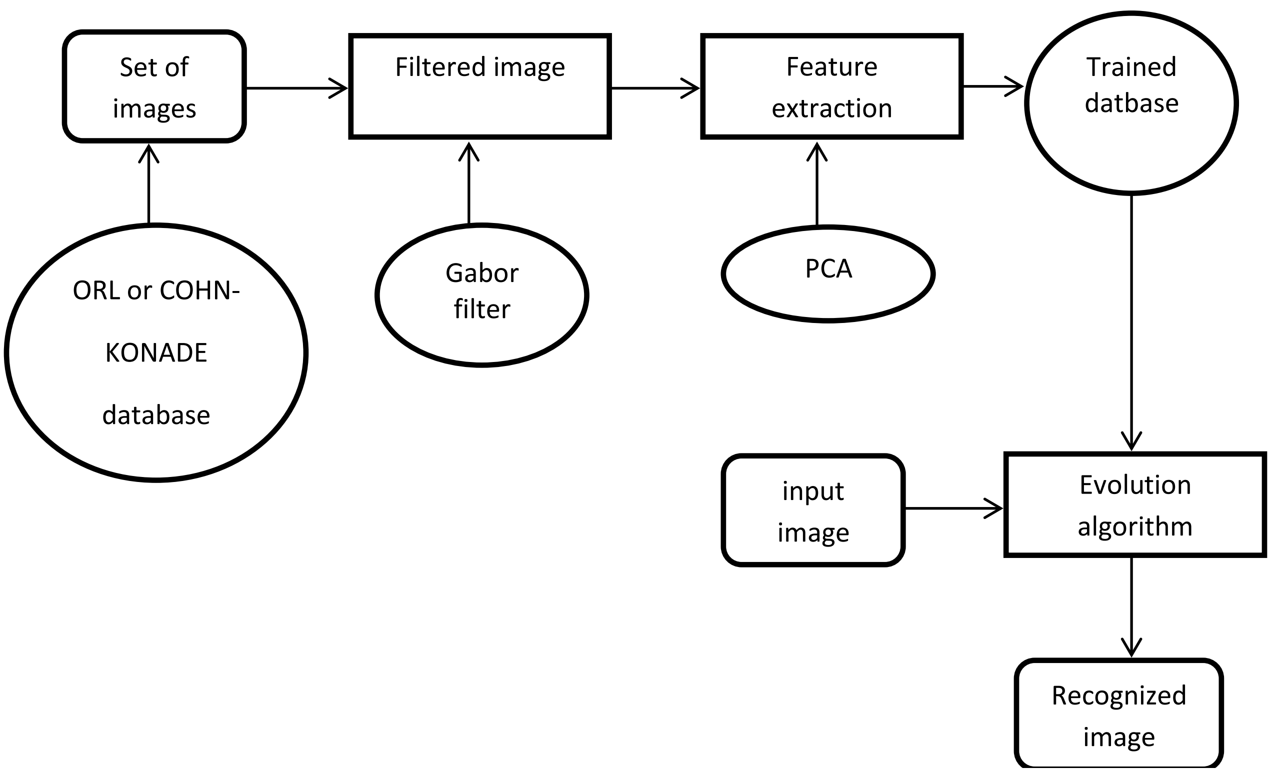


Figure 14 Model for Face recognition

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.

5.2 Databases Taken For Experimentation

We have taken two databases for experimentation. For simulation purpose we have used standard ORL database figure-14 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.(29). 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 [20]. Each image has different characteristic like illumination, facial and pose expression.

Figure 15 ORL database

Figure 16 Cohn-Kanade Database

Cohn Kanade database as shown in fig-15 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.

5.3 Experimental Setup

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

5.3.1 Training Phase

- In case of Exhaustive, PSO, Hybrid ACO/PSO and exhaustive search and Grey Wolf Optimizer :- 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 as there 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 [20]. 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.

- In case of GWO: we divided our database into two parts; train data and test data. In ORL database we used 320 images (8 images out of 10 from all 40 subjects) for training purpose. While in case of cohn-kanade database at training phase 180 images (6 images out 0f 8 from all 30 subjects) for training purpose.

5.3.2 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.

5.4 Results and Analysis

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

5.4.1 On the basis of time taken for recognition

- We have analyzed 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.
- In case of Grey Wolf Optimizer, Face Recognition takes less amount of time. The reason behind this is of using parameters and vectors controlling three possible solution in each iteration as the wolves in their hunting behavior.

	Exhaustive Search	**ACO**	**PSO**	**Hybrid ACO/PSO**	**BBO**	**Extended BBO**	**GWO**
ORL database	**.009012**	**0.03205**	**0.00526**	**0.0237**	**0.00544**	**.008591**	**.00506**
Cohn-Kanade database	**0.00715**	**0.0340**	**0.00264**	**0.0274**	**0.00407**	**0.00858**	**0.0248**

Table 3 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.

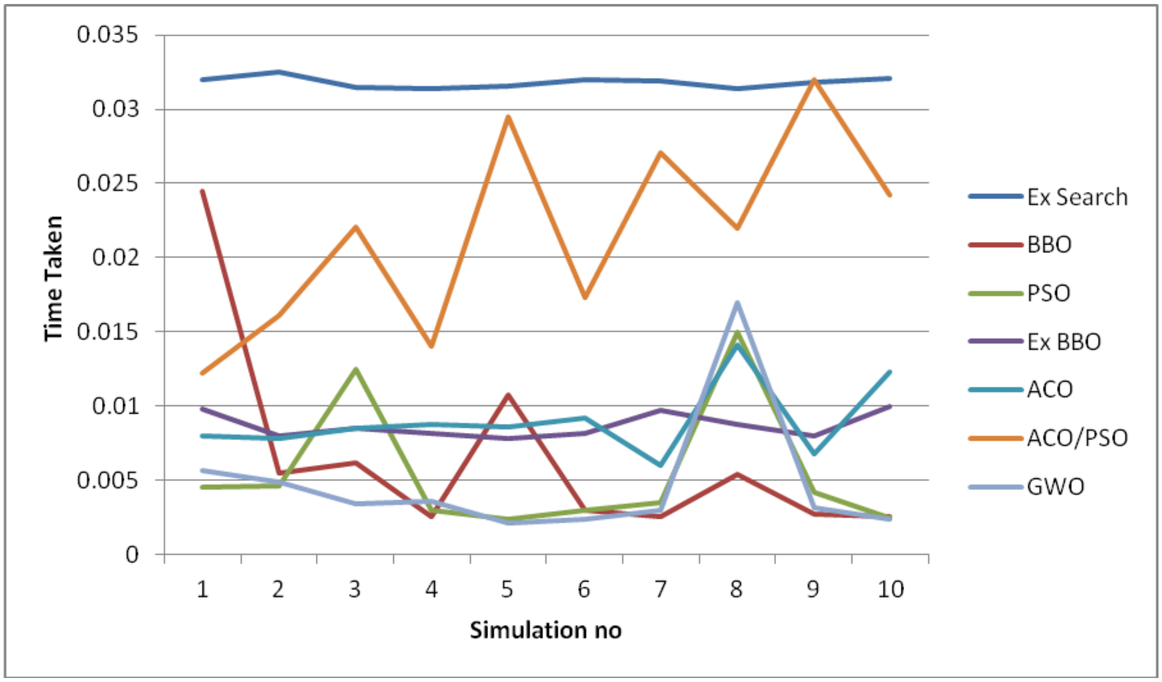


Figure 17 Graph to show simulation results of various algorithms on ORL database

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

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

5.4.2 On the basis of Accuracy

- GWO has the highest accuracy in comparision to all other algorithms except Extended BBO.
- Extended BBO has highest accuracy. Its accuracy is approximately equal to the accuracy of exhaustive search.
- Though PSO takes least time for recognition but it has least accuracy.

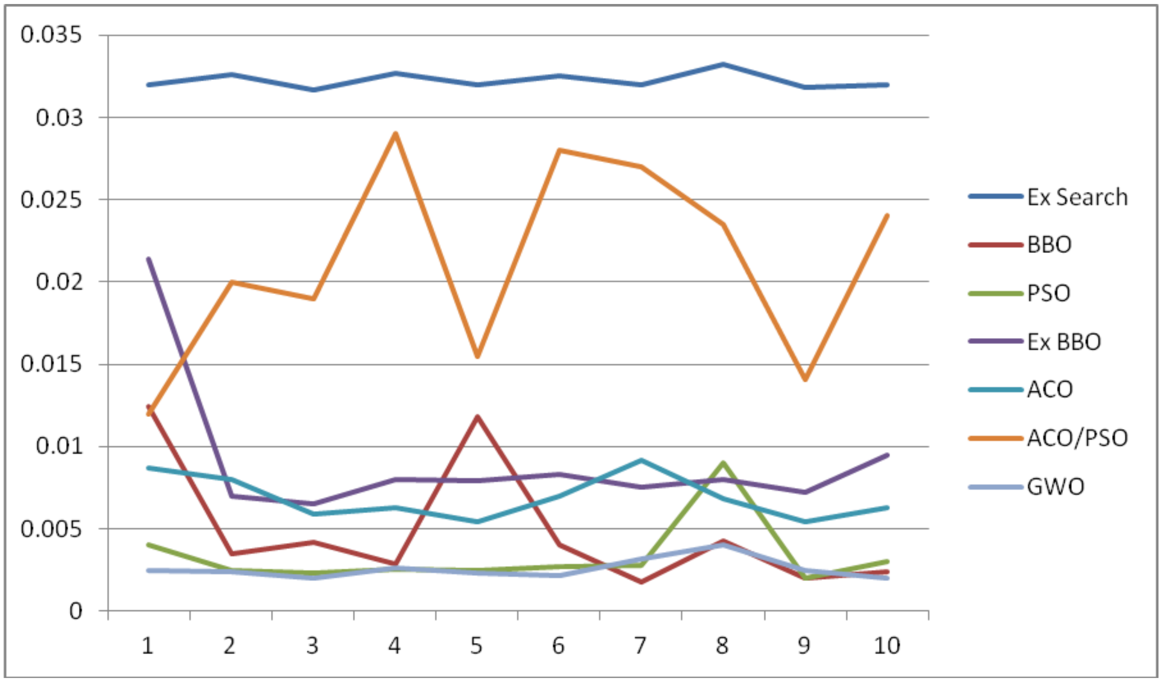


Figure 18 Graph to show simulation results of various algorithms on Cohn-Kanade database.

- 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

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

	Exhaustive Search	**ACO**	**PSO**	**Hybrid ACO/PSO**	**BBO**	**Extended BBO**	**GWO**
ORL database	**90.58**	**83.33**	**50.52**	**81.63**	**84.78**	**92.5**	**91.68**
Cohn-Kanade database	**96.55**	**87.11**	**68.23**	**85.93**	**85**	**98.27**	**97.29**

Table 4 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

6.1 Conclusion

In this thesis we compared various nature inspired algorithms. We compared some advance algorithms like GWO, 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 GWO, PSO, BBO, Extended BBO. We can also use different evolutionary algorithm in training phase and different in recognition phase.

We compared all these algorithms on the basis of running time for recognition and accuracy and we have analyzed that different algorithm is better for different parameter. If we consider time, GWO is best as it takes least time but its accuracy is not good as Extended BBO has.

On the basis of time we have :

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

On the basis of accuracy we have :

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

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.

6.2 Future Work

For feature extraction phase we have used Principal Component Analysis. We can use some other feature 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] Emary, E., et al. "Feature subset selection approach by gray-wolf optimization." Afro-European Conference for Industrial Advancement. Springer International Publishing, 2015.

[2] Mirjalili, Seyedali, Seyed Mohammad Mirjalili, and Andrew Lewis. "Grey wolf optimizer." Advances in Engineering Software 69 (2014): 46-61.

[3] Daya Gupta, Lavika Goel, Kanishka Bansal, "A Novel Approach for Face Recognition Based on Extended Species Abundance Model of Biogeography", In: Proceedings of Second International Conference on Emerging research in Computing, Information, Communication and Applications (ERCICA), Elsevier Publications, Bangalore, 1-2 August, 2014. (presented)

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

[5] Gupta, Daya, Lavika Goel, and Abhishek Abhishek. "An Efficient Biogeography Based Face Recognition Algorithm." 2nd International Conference on Advances in Computer Science and Engineering (CSE 2013). Atlantis Press, 2013.

[6] Desale, Rajenda Pandit, and Sarita V. Verma. "Study and analysis of PCA, DCT & DWT based image fusion techniques." Signal Processing Image Processing & Pattern Recognition (ICSIPR), 2013 International Conference on. IEEE, 2013.

[7] Jigar M. Pandya, Devang Rathod, Jigna J. Jadav," A Survey of Face Recognition approach", International Journal of Engineering Research and Applications (IJERA) ISSN: 2248-9622 www.ijera.com Vol. 3, Issue 1, January -February 2013, pp.632- 635

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

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

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

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

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

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

[14] W. ZHAO,R. CHELLAPPA,P. J. PHILLIPS,A. ROSENFELD. Face Recognition: A Literature Survey ACM Computing Surveys, Vol. 35, No. 4, December 2003, pp. 399–458.

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

[16] Hjelmås, Erik, and Boon Kee Low. "Face detection: A survey." Computer vision and image understanding 83.3 (2001): 236-274

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

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

[19] M. Turk and A. Pentland, ``Eigenfaces for recognition,'' J. Cog. Neuroscience, vol. 3, no. 1, pp. 71-86, 1991

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

[21] M. Kirby and L. Sirovich,``Application of the karhunen-loeve procedure for the characterization of human faces,'' IEEE Pattern Analysis and Machine Intelligence, vol. 12, no. 1, pp. 103-108, 1990.

[22] T. Kohonen, Self-organization and Associative Memory, Springer-Verlag, Berlin, 1989.

[23] KELLY, M. D. 1970. Visual identification of people by computer. Tech. rep. AI-130,

Stanford AI Project, Stanford, CA.

[24] S. Tolba, A.H. El-Baz, and A.A. El-Harby, ” Face Recognition: A Literature Review”, International Journal of Signal Processing 2:2 2006

[25] Jolliffe I.T. Principal Component Analysis, Series: Springer Series in Statistics, 2nd ed., Springer, NY, 2002, XXIX, 487 p. 28 illus. ISBN 978-0-387-95442-4

[26] 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.

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

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

[29] KANADE, T. 1973. Computer recognition of human faces. Birkhauser, Basel, witzerland, and Stuttgart, Germany.

[30] 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.

[31] Emary, E., et al. "Feature subset selection approach by gray-wolf optimization." Afro-European Conference for Industrial Advancement. Springer International Publishing, 2015.

[32] Sharma, Yatin, and Lalit Chandra Saikia. "Automatic generation control of a multi-area ST–Thermal power system using Grey Wolf Optimizer algorithm based classical controllers." International Journal of Electrical Power & Energy Systems 73 (2015): 853-862.

[33] Mustaffa, Zuriani, Yuhanis Yusof, and Siti Sakira Kamaruddin. "Application of Grey Wolf Optimizer for Time Series Forecasting."

[34] Korayem, L., M. Khorsid, and S. S. Kassem. "Using Grey Wolf Algorithm to Solve the

Capacitated Vehicle Routing Problem." IOP Conference Series: Materials Science and Engineering. Vol. 83. No. 1. IOP Publishing, 2015.

[35] Shankar, K., and P. Eswaran. "Sharing a Secret Image with Encapsulated Shares in Visual Cryptography." Procedia Computer Science 70 (2015): 462-468.

[36] Guangpu, Liu, and Gao Yuchun. "The Application of MATLAB in Communication Theory." Procedia Engineering 29 (2012): 321-324.

APPENDIX-A 'ABBREVIATIONS'

PCA – Principal Component Analysis

AS – Ant System

BBO – Biogeography Based Optimization

PSO – Particle Swarm Optimization

ACO – Ant Colony Optimization

PC – Principal Component

HSI – Habitat Suitability Index

ORL – Olivetti Research Laboratory

LDA – Linear Discriminant Analysis

Extended BBO – Extended Species abundance model of biogeography

APPENDIX –B 'MATLAB'

The short form for MATLAB is "matrix laboratory". The software MATLAB 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 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.

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 high levels 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.

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 [36].

List of contents

CHAPTER 1 1
INTRODUCTION 1
1.1 Motivation 2
1.2 Related Work 3
1.3 Problem Statement 4
1.4 Scope of work 5
1.5 Organization of thesis 5
CHAPTER 2 7
LITERATURE REVIEW 7
2.1 Steps of face recognition 8
2.1.1 Face Detection 8
2.1.2 Feature extraction and selection 9
2.1.3 Face recognition 9
2.2 Principle Component Analysis 11
2.2.1 PCA in face recognition 11
2.3 Gabor Filter 12
2.4 Classifier 13
CHAPTER 3 14
GENERIC EVOLUTIONARY TECHNIQUES 14
3.1 Ant Colony Optimization 14
3.1.1 ACO in Face Recognition 15
3.2 Particle Swarm Optimization 17
3.2.1 PSO in Face Recognition 19
3.3 Biogeographical Based Optimization 20
3.3.1 Biogeography Based Optimization (BBO) in Face Recognition 23
3.4 Extended Species Abundance Model Of Biogeography (Extended BBO) 25
3.4.1 Extended BBO in Face Recognition 27
3.5 Hybrid ACO/PSO in Face Recognition 29

CHAPTER 4 31
GWO DESCRIPTION 31
4.1 Inspiration 31
4.2 Mathematical model and algorithm 34
4.2.1 Social hierarchy........ 34
4.2.2 Encircling prey 34
4.2.3 Hunting 35
4.2.4 Search for prey (exploration) 38
4.3 GWO in face recognition 39
4.3.1 Description of algorithm 40
CHAPTER 5 43
MODEL OF FACE RECOGNITION 43
5.1 Face Recognition Model 43
5.2 Databases Taken For Experimentation 44
5.3 Experimental Setup 45
5.3.1 Training Phase 45
5.3.2 Recognition phase 46
5.4 Results and Analysis 46
5.4.1 On the basis of time taken for recognition 46
5.4.2 On the basis of Accuracy 48
CONCLUSION AND FUTURE WORK 50
6.1 Conclusion 50
6.2 Future Work 51
REFERENCES 52
APPENDIX-A ‘ABBREVIATIONS’ 56
APPENDIX –B ‘MATLAB’ 57

LIST OF FIGURES

Figure 1 generic face recognition proces 8
Figure 2 example of recognition process 10
Figure 3 The diagraph 16
Figure 4 species abundance in Single Island 22
Figure 5 Extended Species Abundance Model 26
Figure 6 Steady state probability curve P_k as function of species k 27
Figure 7 Hierarchy of Grey Wolf 31
Figure 8 Hunting Behaviour of Grey Wolves [2]. 33
Figure 9 2D & 3D position vectors and their possible next locations. 35
Figure 10 Position updating in GWO. 36
Figure 11 Attacking prey versus searching for prey. 37
Figure 12 flow chart for Grey Wolf Optimization in Face Recognition 40
Figure 13 Model for Face recognition 43
Figure 14 ORL database 44
Figure 15 Cohn-Kanade Database 45
Figure 16 Graph to show simulation results of various algorithms on ORL database 48
Figure 17 Graph to show simulation results of various algorithms on Cohn-Kanade database. 49

LIST OF TABLES

Table 1 Available Commercial Face Recognition Systems 7
Table 2 adaption of PSO in face recognition 19
Table 3 Time taken by various algorithms on both databases 47
Table 4 Accuracy in both the databases 49