World of Computer Science and Information Technology Journal (WCSIT)
ISSN: 2221-0741
Vol. 2, No. 4, 125-130, 2012

Swap Training: A Genetic Algorithm Based Feature Selection Method Applied on Face Recognition System

Ashkan Parsi
Department of Computer Engineering and IT
Shahrood University of Technology
Shahrood, Iran

Mehrdad Salehi
Computer Aided Medical Procedures
Technische Universität München
Munich, Germany

Ali Doostmohammadi
Department of Computer Engineering and IT
Shahrood University of Technology
Shahrood, Iran

Abstract— This paper presents a new feature selection method by modifying fitness function of genetic algorithm. Our implementation environment is a face recognition system which uses genetic algorithm for feature selection and k-Nearest Neighbor as a classifier together with our proposed Swap Training. In each iteration of genetic algorithm for assessment of one specific chromosome, swaps training switch the training and test data with each other. By using this method, genetic algorithm does not quickly converge to local minimums and final recognition rate will be enhanced. Obtained results from implementing the proposed technique on Yale Face database show performance improvement of genetic algorithm in selecting proper features.

Keywords- face recognition; feature selection; principal component analysis; genetic algorithm; k-Nearest Neighbor classifier.

I. Introduction

We can recognize the other people by set of distinguished features such as eyes color, skin color, height, walking style, wearing style, etc. This ability develops over the early years of childhood and is essential in our social life. Now imagine that you want to help your friend to find and recognize a person in a crowd. Which feature do you select to introduce that person briefly and completely to your friend? Definitely, you will choose distinctive feature of him, for example having a red T-shirt.

Using machines to automate people recognition, especially face recognition is increasing. The problem of face machine recognition has been studied for more than 30 years. It has gained the attention of researchers of several disciplines such as image processing, pattern recognition, computer vision, neural network and computer graphics. Such interest has been motivated by growth of face recognition applications in many areas, including face identification in terminals, user authentication in building access, detecting a known criminal person in metro stations or indexing of and searching for faces in video databases, intelligent computer user interface like authentication in commercial websites, etc.

Face recognition methods can be simply classified into three categories: holistic feature-based matching method, local feature-based matching method and hybrid matching method [1]. In holistic feature-based matching method, the whole face region is used as raw input for the recognition system, like Principal Component Analysis (PCA) projection method [2], Fisher-face method [3], and Nearest Feature Line (NFL) method [4]. An Independent Gabor Features (IGF) method [5] and a kernel Associative Memory (kAM) models-based method [6] were also applied for face recognition.

In local feature-based matching method, the local features such as eyes, nose, and mouth are first extracted and then their locations and local statistics (geometric and/or appearance) are fed into a structural classifier. Geometrical features method [7] and Elastic Bunch Graph Matching (EBGM) method [8] belong to this category.

In hybrid matching method, both holistic and local features are used for recognition purpose. A feature combination scheme for face recognition by fusion of global and local features was presented in [9]. A fully automatic system for face recognition in databases with only a small number of samples was presented by Yan et al [10]. Global and local texture features were extracted and used in the recognition.

Genetic Algorithm (GA) is a kind of optimizers [11-15] which in our research used for selecting the most informative features in conjunction with a classification algorithm [16].

In [17] GA was used to simultaneously editing data set and feature selection for reducing computational time of nearest neighbor classifier. [1] proposed a local feature-based matching method that used GA for feature selection and Artificial Neural Network for classification as well. Many work in face recognition like [1], [17-20] use GA to extract most important feature for improving time and accuracy of their methods.

This paper proposes a new fitness function for GA which has some advantages compared with simple GA. This new fitness function “Swap Training” will generalize the result of feature selection in test phase. We claim that Swap Training is more effective for GA as a fitness function. For more demonstration, obtained results of our new fitness function will compare with the simple GA. In addition, the PCA k-Nearest Neighbor (k-NN) are used for feature extraction and classification respectively.

Generally, this paper is organized as follows. Firstly, relative information about the basic approach for using PCA to extract the face features will be described in section II. Then we will show how GA can select the most informative features among the PCA extracted features in section III and propose a Swap Training technique to improve GA’s efficiency in section IV. The next section (i.e. Section V) introduces the Yale Face database [21] used for testing the proposed approach. Section VI gives the experimental results and finally section VII will present concluding remarks.

II. General Approach

PCA is a useful statistical technique found many applications in fields such as face recognition and image compression [22]. It is a common technique for finding proper patterns in high dimension data it was used to extract the most important feature vectors which will be multiplied to the train dataset [23].

Fig. 1 shows the general approach which is basis for entire methods discussed in current paper. In Fig. 1, new train set that processed by PCA, has then arranged in descending order of features variances. Feature selection can properly help us in reduce the associated computational time and increases the accuracy of classification. Although extracted features of PCA have the most variance, but it doesn’t mean that these features are the more informative ones and can lead to more accuracy in classification process. Therefore, there is a need to select the effective features.

Among the various methods of supervised statistical classification, the Nearest Neighbor rule is one of the best [23]. The k-Nearest Neighbor is an extended algorithm of this rule. It uses the training dataset of each class. A new sample is classified by determining the K distance of the nearest training case. The associated class of each particular sample will be that one which has the greatest repetition in K nearest samples. The Euclidean distance is usually used for calculation of the neighborhood and can be defines as follow:

$$D(Y, X_i) = \sqrt{\sum_{j=1}^{n} (y_j - x_{ij})^2} \quad (1)$$

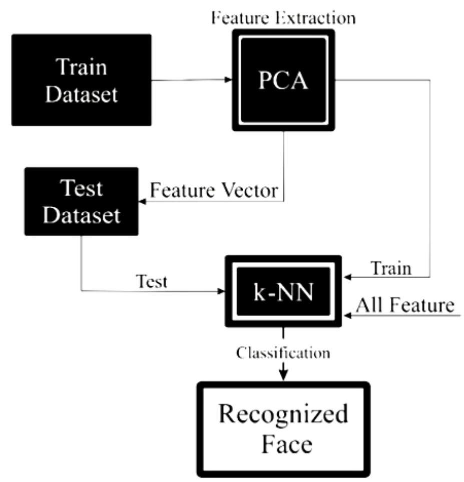

Figure 1. PCA-kNN Approach

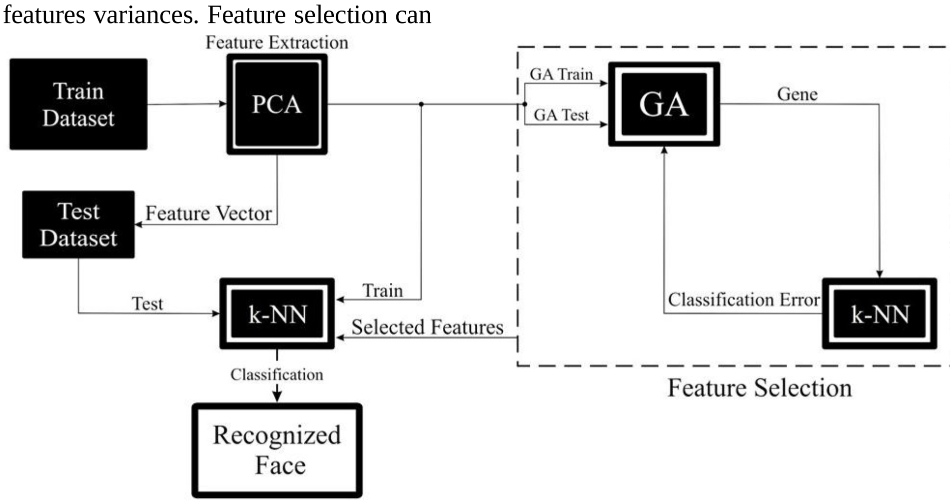

Figure 2. GA-kNN Approach. Using GA for selecting an informative set of features (Selected Features) instead of using All Features for classification by k-NN

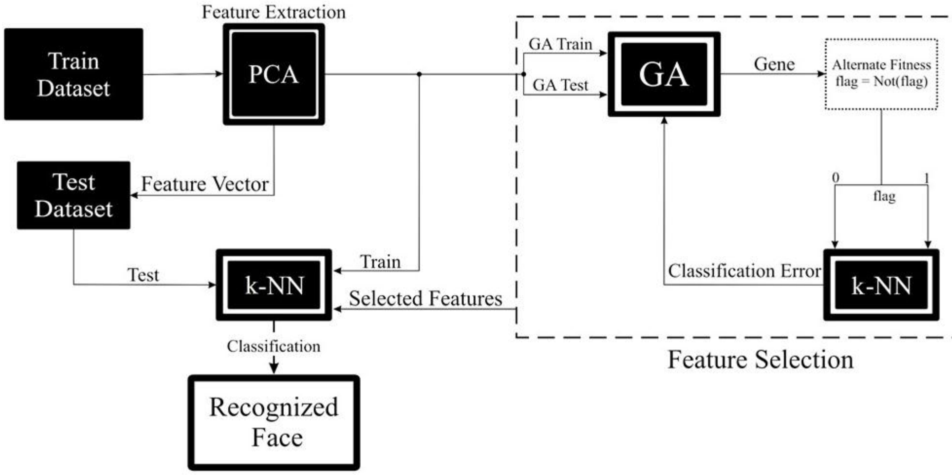

Figure 3. GA-kNN Approach with Swap Training

III. GA-kNN Feature Selection

To improve the efficiency of PCA-kNN approach, the GA has been used to select the features that would increase the performance in both training phase and test phase [18-20], [24, 25]. In first step for using the GA, the training dataset must be partitioned into GA Train dataset and GA Test dataset that is used by GA to learn what features are unique and selects more efficient features.

In each generation, GA generates a set of chromosome and sends each one into the k-NN classifier for calculating the expected fitness. Fitness function is a part of GA which evaluates the accuracy of chromosomes. Each chromosome is the selection mask that determines which features of PCA extracted features should be selected. Finally GA gives the best set of features. Fig. 2 shows this technique.

When selected features used rather than all features in face recognition, the classification accuracy did not increase but the number of features that use to reach to this result is less than PCA-KNN method. This is because the GA tries to select the features in order to fit the GA Train to GA Test as fit as possible and doesn't care about removing the features that would be very useful to recognize new face images correctly.

IV. Proposed Technique: Swap Training

This paper proposes a new technique addressing the GA problem for converging quickly into those features that usually increase the value of GA fitness function, but reduce the recognition accuracy of test samples. The name of our proposed technique is Swap Training. In this novel technique, the alternative fitness function swaps the GA test and train data in each generation and obtained fitness has been evaluated. Fig. 4 shows the algorithm of this function.

A Boolean variable, Flag, determines how the k-NN classifier called by GA test and GA train. It means that the GA_test for the first chromosomes is used as the GA_train of the second one and vice versa. This scenario continues regularly and continuously. Our proposed technique is depicted in Fig. 3.

The proposed technique has improved the performance in manner of accuracy and time in test phase in comparison by simple GA method. This is what we discuss in section VI.

```
SwapFitness ( gene )
{
  if ( flag = 1 )
     fitness = kNN_Fitness( GA_train, GA_test, gene );
  if ( flag = 0 )
     fitness = kNN_Fitness( GA_test, GA_train, gene );
  flag = not( flag );
  return fitness;
}
```

Figure 4. After each iteration, swap fitness function switches the place of GA_train and GA_test which are being used in kNN.

V. Yale Face Database

The Yale Face database [21] containing of 165 persons faces was used as database of this study. As shown in Fig. 5, 11 images are available for each person describing different state of face as follows: center-light (1), glasses (2), happy (3), left-light (4), no glasses (5), normal (6), right-light (7), sad (8), sleepy (9), surprised (10) and wink (11).

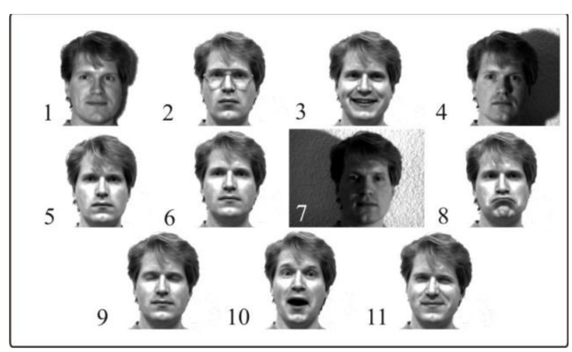

Figure 5. An individual of The Yale Database

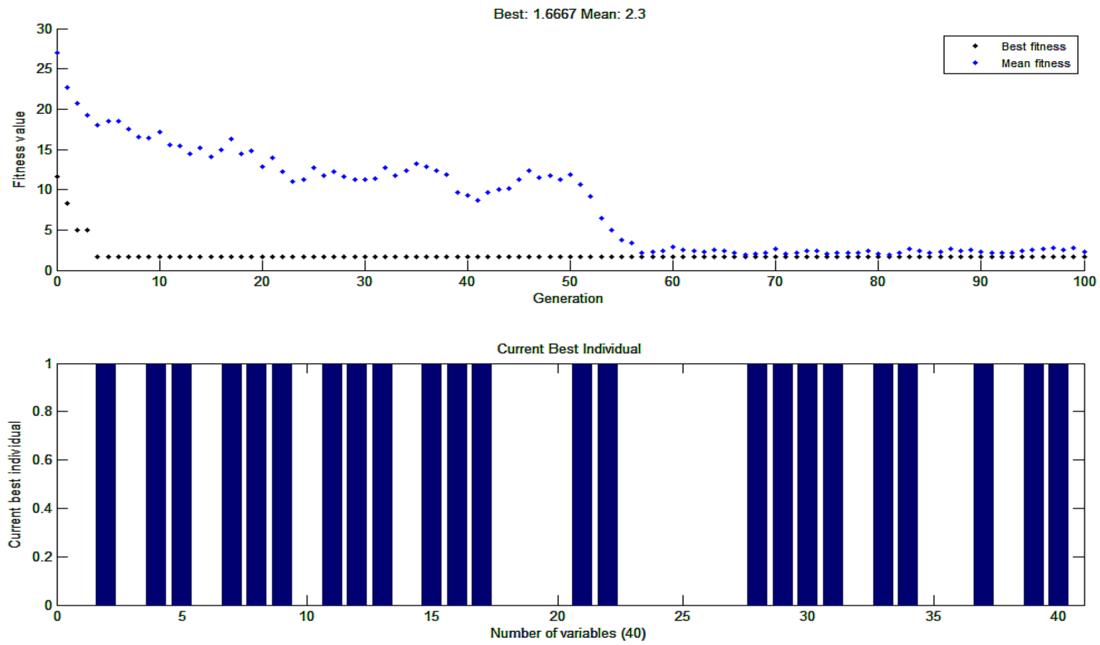

Figure 6. Runtime Results of GA-kNN approach

TABLE I. GA PARAMETERS INITIALIZATION

Genetic Algorithm Parameters	Value
Population Size	50
Type of Population	Bit String
Elite Count	5
Mutation Function	Uniform
Crossover Function	Scattered
Other Parameters	Default Value

TABLE II. GA PARAMETERS CHANGING AT RUNTIME

Genetic Algorithm Parameters	Initialization Value	After 20 Generation	After 50 Generation
Mutation Rate	0.2	0.1	0.01
Crossover Fraction	0.4	0.6	0.8

VI. EXPERIMENTAL RESULTS

In this section, obtained results of using GA-KNN approach with Swap Training in face recognition were shown. Firstly, the initialization value of the GA parameters were introduced, then the way of dataset partitioning into train and test dataset has been described and finally the results presented.

In the experiment, first 40 extracted features of the PCA are selected and then using the proposed feature selection technique, the most detective features extracted to improve the efficiency of the results.

Table I shows initialization values of the GA parameters. Mutation Rate and crossover fraction were being changed at the runtime as shown in Table II. These changes make the GA to quickly converge. So, in the early stages, GA searches into the large solution space and then the solution space is minimized while GA makes advance.

The Yale Face database was partitioned as follows: the train dataset includes pictures number 1, 2, 7 and 8 for GA train, and number 3, 4, 5 and 11 for GA test. The rest of them (i.e. number 6, 9 and 10) were used for test dataset.

Fig. 6 shows the running time results of the GA-kNN approach. As seen in the Figure, simple utilizing of GA-kNN resulted in selecting 23 features within the best fitness of 1.6667 in training phase, while proposed method (i.e. Swap Training) could reach to the similar fitness using 18 features. Fig. 7 can provide a better insight into what we have already said. Also the difference between best fitness and mean fitness have been increased when Swap Training is used. This means that searching process will be done with higher variance in our space and the chromosomes converge to the best ones. Hence, selection of the best features would improve the face recognition accuracy.

In Fig. 8 the comparison results of GA-kNN and Swap training is shown. It shows that there is no improvement in GA-kNN method without Swap training compared with PCA. However, GA-kNN method with Swap training can improve the classification accuracy about 2.2%.

VII. CONCLUSION AND FUTURE WORKS

In this paper, a new feature selection technique for face recognition has been presented. Among the PCA extracted features, the most proper ones should be selected to enhance the performance of classification. Although GA considered as one the best optimization methods, defining an appropriate and global fitness function for the feature selection has a high impact on its performance.

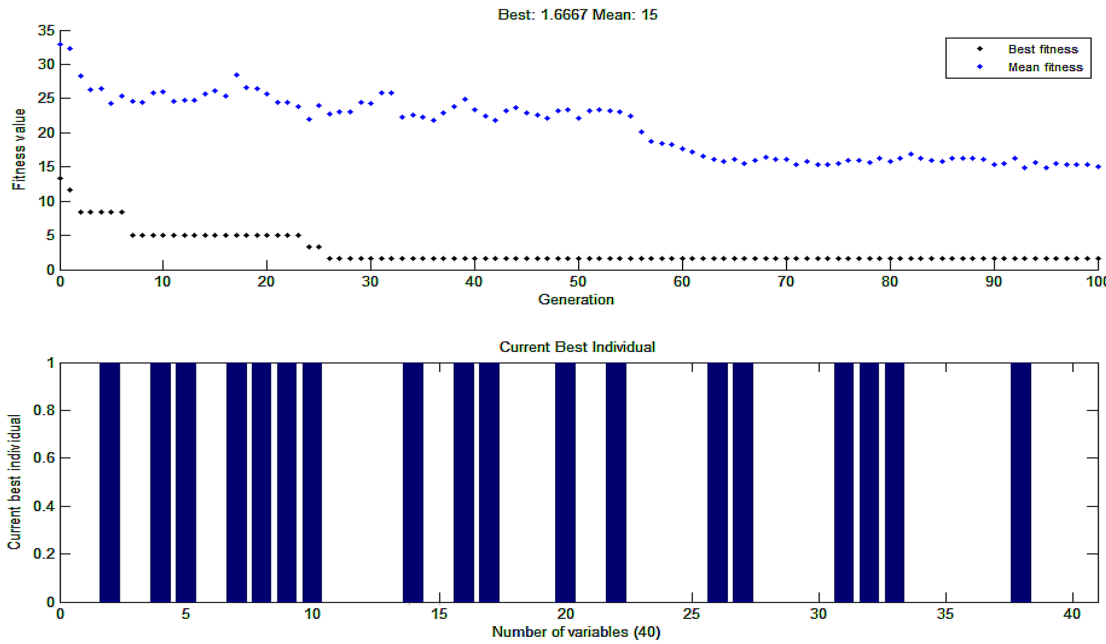

Figure 7. Runtime Results of GA-kNN Approach with Swap Training

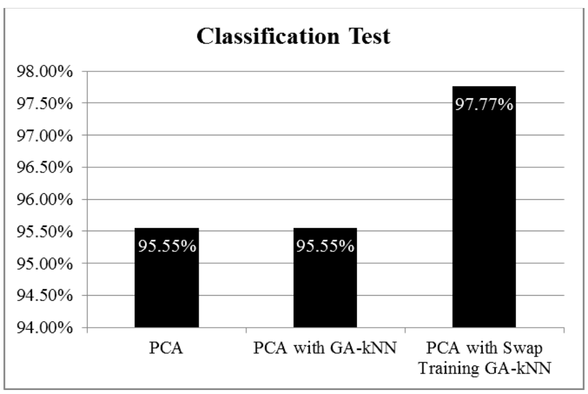

Figure 8. Comparision Results

The associated problem of simple GA fitness function was the quick convergence into uninformative feature sets. In the proposed techniques named Swap Training, a new fitness function introduced to deal with this challenge. Swap Training can properly improve the GA-kNN approach by 2.22% of recognition rate on Yale Face database.

REFERENCES

[1] X. Fan and B. Verma, “Selection and fusion of facial features for face recognition,” Expert Systems with Applications: an International Journal, vol. 36, pp. 7157-7169, April. 2009.

[2] M. Turk, A. Pentland, “Eigenfaces for recognition,” Journal of Cognitive Neuroscience, vol. 3, pp. 71-86, 1991.

[3] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, “Eigenfaces vs. fisherfaces: Recognition using class specific linear projection,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 19, pp. 711-720, 1997.

[4] S. Z. Li and J. Lu, “Face Recognition Using the Nearest Feature Line Method,” IEEE Transactions on Neural Networks, vol. 10, pp. 439-443, 1999.

[5] Ch. Liu and H. Wechsler, “Independent Component Analysis of Gabor Features for Face Recognition,” IEEE Transactions on Neural Networks, vol. 14, pp. 919-928, 2003.

[6] B. Zhang, H. Zhang, and S. Ge, “Face recognition by applying wavelet sub band representation and kernel associative memory,” IEEE Transactions on Neural Networks, vol. 15, pp. 166–177, 2004.

[7] R. Brunelli and T. Poggio, “Face recognition through geometrical features,” In Proceedings of European Conference on Computer Vision, pp. 792–800, 1992.

[8] L. Wiskott, J. M. Fellous, N. Krueuger, and C. Von Der Malsburg, “Face Recognition by Elastic Bunch Graph Matching,” Chapter 11 in Intelligent Biometric Techniques in Fingerprint and Face Recognition, L.C. Jain et al., CRC Press, pp. 355-396, 1999.

[9] Y. Fang, T. Tan, and Y. Wang, “Fusion of global and local features for face verification,” IEEE International Conference on Pattern Recognition, pp. 382–385, 2002.

[10] S. Yan, X. He, Y. Hu, H. Zhang, M. Li, and Q. Cheng, “Bayesian shape localization for face recognition using global and local textures,” IEEE Transactions on Circuits and Systems for Video Technology, vol. 14, No. 1, pp. 102–113, 2004.

[11] K. A. Jong and W. M. Spears, “Using genetic algorithm to solve NP-complete problem,” George Mason University, 1989.

[12] D. M. Mukhopadhyay, M. O. Balitanas, A. A. Farkhod, S. H. Jeon, and D. Bhattacharyya, “Genetic algorithm: A tutorial review,” International Journal of Grid and Distributed Computing, vol. 2, no.3, pp. 25-32, 2009.

[13] S. N. Sivanandam and S. N. Deepa, Introduction to Genetic Algorithms, Springer Berlin Heidelberg New York, 2008.

[14] H. Vafaie and K. De Jong, “Genetic algorithms as a tool for restructuring feature space representations,” Tools with Artificial Intelligence, Seventh International Conference, 1995.

[15] D. Whitley, “A genetic algorithm tutorial. Statistic and Computing,” Springer, vol. 4, no. 2, pp. 65-85, 1994.

[16] F. Hussein, N. Kharma, and R. Ward, “Genetic algorithms for feature selection and weighting, a review and study,” Document Analysis and Recognition, Sixth International Conference, 2001.

[17] L. I. Kuncheva and L. C. Jain, "Nearest neighbor classifier: Simultaneous editing and feature selection," Pattern Recognition Letters, vol. 20, pp. 1149-1156, 1999.

[18] M. T. Harandi, M. N. Ahmadabadi, B. N. Araabi, and C. Lucas, "Feature selection using genetic algorithm and it's applications to face recognition," Cybernetics and Intelligent Systems, IEEE Conference, 2004.

[19] J. Jarmulak and S. Craw, "Genetic algorithms for feature selection and weighting," Appears in Proceedings of the IJCAI'99 workshop on Automating the Construction of Case Based Reasoners, 1999.

[20] D. Mohamad, "Multi local feature selection using genetic algorithm for face identification," International Journal of Image Processing, vol 1, pp. 1-10, 2008.

[21] The Yale Face Database, http://cvc.yale.edu/projects/yalefaces/yalefaces.html.

[22] H. Moon and P. J. Phiilps, "Computational and performance aspect of PCA-based face recognition algorithm," Perception, vol. 30, pp. 303-321, 2001.

[23] K. Fukunaga, Introduction to Statistical Pattern Recognition, 2nd ed., Academic Press, New York, 1990.

[24] S. Venkatesan and S. Srinivasa Rao Madane, "Face recognition system with genetic algorithm and ANT colony optimization," Appears in Proceedings of the IJCAI'99 workshop on Automating the Construction of Case Based Reasoners, 1999.

[25] G.G. Yen and N. Nithianandan, "Facial feature extraction using genetic algorithm," Evolutiory Computation CEC '02,2002.