

Genetic Algorithm based Feature Subset Selection in Face Detection

Pratibha P. Chavan¹, Dr. M Murugan², Pramod U. Chavan³

¹ Research Scholar JJTU, Rajasthan and Assistant Prof. of Electronics & Telecommunication Engineering Dept., K J College of Engineering & Management Research, Pune-, India.

² Vice-Principal & Head of Electronics & Communication Engineering Department, SRM Valliammai Engineering College, Kattankulathur, Chennai-, India.

³ Assistant Professor, Department of Electronics & Telecommunication Engineering, Vishwakarma Institute of Information Technology, Pune-, India.

ABSTRACT: Genetic Algorithm is used in this paper, which presents the designing an optimized tool for face detection application with idea of "Survival of the fittest". If higher level of adaption can be achieved, existing systems can perform their functions longer & better. Popular method in feature extraction is principal component analysis (PCA), for feature subset selection used Genetic algorithm (GA) & support vector machines (SVM) used for classification. The goal is to search PCA space using GA & select a subset of eigen vectors encoding important information about the target concept of interest.

Keywords- Face Detection, Feature extraction, Genetic Algorithm, Principal Component Analysis, Support Vector Machine

I. INTRODUCTION

Referring to facial images, its detection in an image is a problem that requires a precise research due to its high complexity. The usage of these images is very wide, such is the security systems case, where it can be used for remote conference, search in database, people identification, and so on [1]. The increase of terrorist attacks motivates countless works in this segment. Security cameras are placed in Main roads, Highways, supermarkets, shopping centers, parks, buses, and so on, trying to help the identification of blameworthy people, besides reducing crimes. Unfortunately, until now, security systems need human supervise, which causes a significant number of failures. The approach is based on the Genetic algorithm & eigen-face technique.

II. FACE DETECTION

The problem of visual Face detection is easy for people but has proved to be very difficult for computer. Face detection system is unique method. The purpose of the face detection is to search and orient faces in images in complex background [2]. The frequently used biometric features include face, fingerprint, voice, and iris recognition. The fingerprint recognition is the most popular adopted in our daily lives. However, the sweat and the dust may reduce the accuracy sometimes. In face detection system, it is not necessary to have physical contact with the machine and the image can be captured naturally by using a video camera. This makes face detection a very convenient biometric identification approach. Face detection uses in Surveillance, Recognition, personal security and so many other applications [3].

III. GENETIC ALGORITHM

John Holland, his colleagues and his students at the University of Michigan have been developed Genetic Algorithm in the 1960s [4]. Features for self-repair, self-guidance and reproduction are the rule in biological systems, whereas they barely exist in most sophisticated artificial systems. The goal of GA is to build an artificial genetic system with natural characteristic and makes each species eliminate through competition, that only good adaptable species can survive and propagate to produce the coming generation [5]. Mitchell used three basic operators in GA which are Selection, Crossover and Mutation [6].

1. Selection: Selection or reproduction is usually the first operator applied to population. This operator selects good string (high fitness value) in a population and forms a mating pool. This operator simulates the "Survival of the fittest" and the idea is to give performance to better individuals.

2. Crossover: In the crossover operator, two strings are picked from the mating pool at random and some portions are of the strings are exchanged between the strings by single point crossover i.e. exchanging information among strings of mating pool creates new strings.

3. Mutation: The mutation operator plays a secondary role. This operator selects a random bit with a small probability from string and changes a bit to its compliment.

IV. FEATURE EXTRACTION USING PRINCIPLE COMPONENT ANALYSIS

The purpose of PCA is to reduce the large dimensionality of the data space (observed variables) to the smaller intrinsic dimensionality of feature space (independent variables), which are needed to describe the data economically. This is the case when there is a strong correlation between observed variables. Take 2D images (Scaled to 19 X 19 pixels) of n sample faces and non faces from database. Each pixel can take real values between 0 and 255. Each image is considered to be a single sample from a random vector. Feature extraction means presented data in lower dimensions using Principle Component Analysis (PCA) [7]. It is used in selection of the feature vectors related to the biggest eigen value. Find the mean image m and subtract it from each image. $Z_i = x_i - m$ and stored into the matrix whose columns are the mean-subtracted sample images.

$$P = \begin{matrix} \uparrow & \uparrow & \uparrow \\ Z_1 & Z_2 & \dots & Z_n \\ \downarrow & \downarrow & \downarrow \end{matrix} \quad (1)$$

Estimate the covariance matrix: $\sum \lambda_i = \frac{1}{n} P P^T$ (2)

Since Σ is very large (361 X 361), computing its eigen vector will be very expensive. Instead, we can compute λ_i , the eigen vectors of $P^T P$ (Transpose Trick), an 19×19 matrix. Then μ_i can be computed from λ_i as follows:

$$\mu_i = \sum_{j=1}^n \lambda_{ij} Z_j \quad j = 1, \dots, n \quad (3)$$

Usually, we only need to keep a smaller number of eigenvectors nk corresponding to the largest eigen - values. Given a new image Γ , we subtract the mean $\Phi = \Gamma - m$ and compute the projection:

$$\phi = \sum_{i=1}^{Rk} w_i \mu_i \quad w_i = \mu_i^T \Gamma \quad (4)$$

V. CLASSIFICATION USING SUPPORT VECTOR MACHINE

Support vector machines (SVM) were first suggested by Vapnik (1995). SVM classifies data with different class labels by determining a set of support vectors that are members of the set of training inputs that outline a hyper plane in the feature space [8]. When using SVM, two problems are confronted 1st problem is how to choose the optimal input feature subset for SVM, and 2nd is how to set the best kernel parameters. These two problems are crucial, because the feature subset choice influences the appropriate kernel parameters and vice versa (Frohlich and Chapelle, 2003). In addition to the feature selection, proper parameters setting can improve the SVM classification accuracy [9]. The parameters that should be optimized include penalty parameter C and the kernel function parameters such as the gamma (g) for the radial basis function (RBF) kernel. To design a SVM, one must choose a kernel function, set the kernel parameters and determine a soft margin constant C (penalty parameter). The classification problem can be restricted to consideration of the two-class problem without loss of generality.

The goal is to separate the two classes by a function. The SVM finds the hyper plane leaving the largest possible fraction of points of the same class on the same side, while maximizing the distance of either class from the hyper plane. Assuming there are s examples from two classes: (x1, y1) (x2, y2)... ... (Xs, ys)

$$x_i = \in R^n, y_i = \in \{-1, +1\} \quad (5)$$

Finding the optimal hyper-plane implies solving a constrained optimization problem using quadratic programming. The discriminate hyper plane is defined as

$$F(x) = \sum_{i=1}^s y_i a_i k(x, x_i) + b \quad (6)$$

Where $k(x; x_i)$ is a kernel function and the sign of $f(x)$ indicates the membership of x. Each function satisfying the Mercer condition can be used here. The Gaussian radial basis kernel is given by

$$K(x, x_i) = \exp\left(-\frac{\|x-x_i\|^2}{2\sigma^2}\right) \quad (7)$$

The Gaussian kernel is used in this study. Experiments have shown that the Gaussian kernel outperforms other kernels in the context of our applications.

VI. FACE DETECTION USING GA

The goal of feature subset selection is to use less features to achieve the same or better performance [10]. Therefore, the fitness evaluation contains two terms: 1) Accuracy and 2) Number of features selected. Chromosome contains n genes evaluated by 0 or 1. If eigen vector is chosen represented by 1 and 0 means the eigen vector is not chosen. Take n chromosomes. The initial population after each selection will be constant. Uniform crossover is used here and the mutation probability used in all of experiments is 0.01. Each feature subset contains a certain number of eigen vectors. If two subsets achieve the same performance, while containing different numbers of eigen vectors, the subset with fewer eigen vectors is preferred. Between

accuracy and feature subset size, accuracy is our major concern. We used the fitness function below to combine the two terms:

$$\text{Fitness} = 10^4 \text{ Accuracy} + \text{Zeros} \quad (8)$$

Classification accuracy, the number of selected features, and the feature cost are the three criteria used to design a fitness function. Thus, for the individual (chromosome) with high classification accuracy, a small number of features, and low total feature cost produce a high fitness value. We solve the multiple criteria problem by creating a single objective fitness function that combines the three goals into one. The chromosome with high fitness value has high probability to be preserved to the next generation, so user should appropriately define these settings according to his requirements. Accuracy corresponds to the classification accuracy on a validation set for a particular subset of eigen vectors, and Zeros corresponds to the number of eigen vectors not selected (i.e., zeros in the chromosome). Based on the weights that we have assigned to each term, the Accuracy term dominates the fitness value. This implies that individuals with higher accuracy will outweigh individuals with lower accuracy, no matter how many features they contain. Overall, the higher the accuracy, gives the higher the fitness. Also, the fewer the number of features, gives the higher the fitness [1]. Accuracy using the binary target data sets can be demonstrated by the positive hit rate (sensitivity), the negative hit rate (specificity), and the overall hit rate. For the multiple class data sets, the accuracy is demonstrated only by the average hit rate. Some cases with the ‘positive’ class (with disease) correctly classified as positive (TP = True Positive fraction), however, some cases with the ‘positive’ class will be classified negative (FN = False Negative fraction). Conversely, some cases with the ‘negative’ class (without the disease) will be correctly classified as negative (TN= True Negative fraction), while some cases with the ‘negative’ class will be classified as positive (FP= False Positive fraction).

TP and FP are the two important evaluation performances for classifiers. Sensitivity and specificity describe how well the classifier discriminates between case with positive and with negative class (with and without disease). Sensitivity is the proportion of cases with positive class that are classified as positive (true positive rate, expressed as a percentage). Specificity is the proportion of cases with the negative class, classified as negative (true negative rate, expressed as Percentage). In probability notation for

$$\text{Sensitivity} = \frac{TP}{(TP+FN)} \quad (9) \qquad \text{Specificity} = \frac{TN}{(TN+FP)} \quad (10)$$

$$\text{Overall hit rate} = \frac{(TP+TN)}{(TN+FP+FN+FP)} \quad (11)$$

The SVM accuracy of the fitness in function is measured by sensitivity * Specificity for the data sets with two classes (Positive or negative), and by the overall hit rate for the data sets with multiple classes. Targets are shown in following table no 1.

	Target (or disease)		
		+	-
Predicted (or test)	+	True Positive (TP)	False Positive (FP)
	-	False Negative (FN)	True Negative (TN)

Table No. 1: (2 X 2) contingency table

VII. EXPERIMENTATION

Method	Error % of test data	Fitness
60 eigen vector which corresponding to biggest eigen-value	31.25%	6815
50 eigen vector which corresponding to biggest eigen-value	37.50%	6300
Genetic Algorithm	12%	6850

Table No. 2 Comparison of error percentage for all method

TEST 1:

Enter the no. of non-face images = 50 Enter the no. of face images = 50 Enter the of population =200
 Enter the no of Generation =300

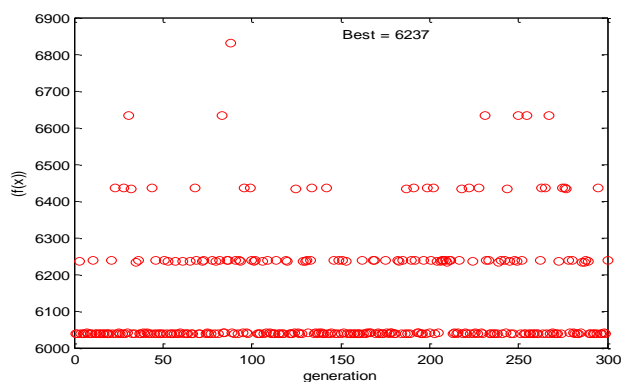


Fig.1. Fitness calculation



Fig 2 Image (size 19 X 19) No. 12 from database
The given image contains the **face**

TEST 2:

Enter the no. of non-face images= 50 Enter the no. of face images = 50 Enter the of population =16
Enter the no of Generation =50

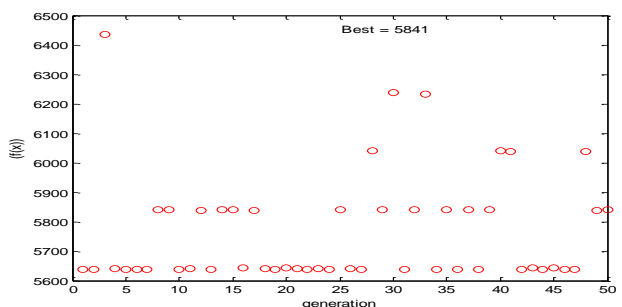


Fig 3 Fitness calculation



Fig 4 Image (size 19 X 19) No. 27 from database
The given image contains the **face**

VIII. DISCUSSION & CONCLUSION

Genetic algorithms is based on natural characteristic and face detection is unique biometric identification system. GA are best employed on problems where it is difficult or impossible to test for optimality and they are efficient for face detection. In this paper we conclude that the ability of intelligent system in search space. Genetic algorithm is better method for finding the best subset. The Principle Component Analysis (PCA) can reduce the dimension of input data space by discarding the components having relatively small eigen-values. Support vector machine classified into face and non face images by providing it training set with selected feature subset and testing set with selected feature subset. We conclude that PCA is a fast & accurate method & Genetic algorithm used for feature subset selection. GA gives better fitness & less error.

REFERENCES

1. Mahendra Pratap Panigrahy, Neeraj Kumar "Face Recognition using Genetic Algorithm and Neural Networks" by IJCA Journal Vol 55 - Number 4, 2012
2. Roshan Bhaiswar, Pravin Kshirasagar, Ambarish A. Salodkar, "A Novel Approach for Face Detection in Complex Background Using Genetic Algorithm", International Journal of Engineering and Innovative Technology (IJEIT) Volume 1, Issue 3, March 2012
3. M.A.Dabbah, W. L. Woo, and S. S. Dlay, "Secure Authentication for Face Recognition," In Proc. of IEEE Symposium on Computational Intelligence in Image and Signal Processing, Apr. 2007. USA, pp. 121 - 126.
4. David E. Goldberg, "Genetic algorithms in search optimization and machine learning," Addison-Wesley 1989.

5. Pramod U Chavan, Pratibha P Chavan, "Codebook Optimization in Vector Quantization using Genetic Algorithm" ICCEE '09 Proceedings of the 2009 Second International Conference on Computer and Electrical Engineering - Volume 01 (Published in the conference proceeding by IEEE Computer Society CPS, listed in IEEE Xplore and CSDL, indexed by INSPEC, Thomson ISI proceeding, and Ei Compendex.)
6. Mealane Mitchell, "An introduction to Genetic algorithm," The MIT Press 1996.
7. B. Mener and F. Muller, "Face Detection in Color Images Using Principal Components Analysis", Image Processing and its Applications, IEEE, 1999.
8. Peng wang, "Multi-view face detection under complex scene based on combined SVMs" Pattern Recognition, 2004. ICPR-2004. Proceedings of the 17th International Conference.
9. Zalhan Mohd Zin, Marzuki Khalid & Rubiyah Yusof, "Evolutionary Feature Selections for Face Detection System." IEEE 2008.
10. Z. Suna, G. Bebisa, R. Millerb, " Object detection using feature subset selection", Elsevier, The journal of Pattern Recognition, March 2004.