Expert Systems with Applications 36 (2009) 7157–7169

Contents lists available at ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

Selection and fusion of facial features for face recognition

Xiaolong Fan, Brijesh Verma *

School of Computing Sciences, Faculty of Business and Informatics, Central Queensland University, Rockhampton, QLD 4701, Australia

ARTICLE INFO

Keywords:
Face recognition
Neural networks
Evolutionary algorithms
Pattern recognition

ABSTRACT

This paper proposes and investigates a facial feature selection and fusion technique for improving the classification accuracy of face recognition systems. The proposed technique is novel in terms of feature selection and fusion processes. It incorporates neural networks and genetic algorithms for the selection and classification of facial features. The proposed technique is evaluated by using the separate facial region features and the combined features. The combined features outperform the separate facial region features in the experimental investigation. A comprehensive comparison with other existing face recognition techniques on FERET benchmark database is included in this paper. The proposed technique has produced 94% classification accuracy, which is a significant improvement and best classification accuracy among the published results in the literature.

1. Introduction

1.1. Background

Face recognition is one of the most remarkable capabilities of human beings. It develops over the early years of childhood, and is important for several aspects of our social life. Human beings can remember hundreds or even thousands of faces in their whole life and can easily identify a familiar face in different perspective variations, such as illumination variations, age variations, and pose variations. Face recognition together with other abilities, such as estimating the expression of people with whom we interact, has played an important role in the course of evolution.

The problem of machine recognition of faces has been studied for more than 30 years. It has attracted research interest from several disciplines such as image processing, pattern recognition, computer vision, neural networks and computer graphics. Such interest has been motivated by the growth of Face Recognition Technology (FRT) used in the applications in many areas, including face identification in law enforcement and forensics, user authentication in building access or automatic teller machines, indexing of, and searching for, faces in video databases, intelligent computer user interfaces, etc. After the September 11, 2001, terrorist attacks, FRT has been gaining more interest due to its significant involvement in anti-terror activities. FRT numerously used in commercial and law enforcement applications poses a wide range of technical challenges and requires an equally wide range of techniques from different disciplines.

A general statement of the problem of machine recognition of faces can be described as follows: Given still or video images of a scene, identify or verify one or more persons in the scene using a stored database of faces. Available collateral information such as race, gender, age, facial expression or speech may be used in narrowing the search. The solution to the problem involves face detection, feature extraction from face region, face verification or recognition. Face detection refers to the determination of the exact position and size of a human face from cluttered scenes. Feature extraction refers to obtaining the features that can be fed into a face classification system. Face recognition refers to comparing an input face against models of faces that are stored in a database of known faces and then indicating if a match is found. Face verification refers to confirming or rejecting the claimed identity of the input face.

Although human beings seem to recognize a face in cluttered scenes with relative ease, machine recognition is much more difficult for a variety of reasons. Firstly, different faces may appear very similar, i.e. every face contains two eyes, two ears, one nose and one mouth, thereby necessitating an exacting discriminant task. Secondly, different views of the same face may appear quite different due to imaging constraints, such as changes in illumination and variability in facial expressions, and due to the presence of personal accessories, such as glasses, beards, and hats. Finally, when the face undergoes rotations out of the imaging plane, a large amount of detailed facial structure may be occluded. Therefore, until now in many implementations of face recognition algorithms, the face images are obtained in a constrained environment with controlled illumination, minimal occlusions of facial structures, uncluttered background, and so on. Face recognition in an unconstrained environment is still a quite challenging task.

* Corresponding author.
E-mail addresses: x.fan@cqu.edu.au (X. Fan), b.verma@cqu.edu.au (B. Verma).

0957-4174/$ - see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.eswa.2008.08.052

1.2. Literature review

In the last decade, face recognition has become one of most active research areas of pattern recognition. The most existing face recognition methods can be simply classified into three categories: holistic feature-based matching method, local feature-based matching method and hybrid matching method (Chellappa, Wilson, & Sirohey, 1995). In holistic feature-based matching method, the whole face region is used as raw input to the recognition system, like Principal Component Analysis (PCA) projection method (Turk & Pentland, 1991), Fisher-face method (Belhumeur, Hespanha, & Kriegman, 1997) and Nearest Feature Line (NFL) method (Li & Lu, 1999). Recently, an Independent Gabor Features (IGF) method (Liu & Wechsler, 2003) and a kernel Associative Memory (kAM) models-based method (Zhang, Zhang, & Ge, 2004) were also applied to 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 (Brunelli & Poggio, 1992) and Elastic Bunch Graph Matching (EBGM) method (Wiskott, Fellous, & Malsburg, 1997) belong to this category.

In hybrid matching method, both holistic and local features are used for the recognition. A feature combination scheme for face recognition by fusion of global and local features is presented in Fang, Tan, and Wang (2002). A fully automatic system for face recognition in databases with only a small number of samples is presented in Yan et al. (2004). Global and local texture features are extracted and used in the recognition.

Genetic Algorithms (GAs) can be used to select an optimal feature set for pattern classification problems. Some researchers have used GAs for face recognition. In Bala, Huang, Vafaie, DeJong, and Wechsler (1995), GA–ID3 (decision tree learning) method is proposed to find an optimal subset of discriminatory features for pattern classification. GAs were used to search the possible optimal subset of extracted features. ID3 was used to produce a decision tree based on a subset selected by GAs. The GA–ID3 method was experimented to recognize visual concepts in satellite and face images. The results showed a significant improvement in classification performance and a good reduction in feature set dimension. In Liu, Tang, Lu, and Ma (2004), a kernel scatter-difference-based discriminant analysis for face recognition is presented. In Sun and Yin (2005), a genetic algorithm was used to select features for 3D face recognition. The method presented in Sun and Yin (2005), tries to optimise features by capturing good features which can minimize the inner-class distance and maximize the intra-class distance. In Liu and Wechsler (2000), an evolutionary pursuit (EP) based on GAs has been applied to face recognition. The idea in EP is to search for face basis through the rotated axes defined in PCA space. The overall classification rate obtained by the existing techniques is unsatisfactory; therefore there is a need for a better feature selection and fusion technique which could improve the overall classification accuracy for face recognition.

In this paper, a novel feature selection and fusion technique for face recognition is presented. GA for feature selection and Artificial Neural Network (ANN) for classification were incorporated into the proposed technique. The proposed technique has been tested on a separate feature set from each facial region and compared with the combined feature set. A large set of dataset from the FERET benchmark database (Phillips, Wechsler, Huang, & Rauss, 1998) is used for testing. The main research questions are (1) How to select the most significant facial features and combine them to improve an overall classification rate of face recognition systems? (2) What is the best combination of these features to a specific classifier? The original contributions of the research presented in this paper are as follows: (1) Identification of local facial regions by using distance threshold method based on center coordinate information of each facial region. The facial features are extracted from each facial region. (2) A Genetic Algorithms (GAs)-based approach for facial feature selection. The significant areas inside each facial region are located using this approach. (3) An Artificial Neural Network (ANN)-based approach for facial feature classification. The selected facial features from GA approach are passed to ANN for final classification. The classification error is passed back to GA to calculate the fitness of each individual. (4) A combined technique for face recognition. The proposed approach is tested on the separate feature set from each facial region and the combined feature set. The FERET benchmark database is adopted to evaluate and compare the proposed approach. A comprehensive comparison of the proposed technique with other existing face recognition approaches has been conducted.

2. Proposed technique

This section describes the proposed feature selection and fusion technique for face recognition. Section 2.1 provides an overview of the proposed methodology. Section 2.2 introduces the distance threshold method that is used to locate facial regions. The average grey level value features are discussed in Section 2.3. Section 2.4 describes PCA features. The details of incorporating GAs and neural networks for feature selection and classification are discussed in Section 2.5.

2.1. Overview

The goal of the proposed technique is to select the most significant facial features effectively and find the best combination of these features for the classifier. The proposed technique aims to locate the significant areas in facial regions from which the significant features are extracted. Facial regions refer to the separate regions in the face that contain one local organ, such as left eye region, right eye region, nose region and mouth region. These facial regions contain the most discriminant facial characteristics on human faces. The facial regions are the basis for the local feature-based feature extraction techniques. Even on these discriminant facial regions, some areas inside may be more important than the other areas in a recognition task. By locating the most significant areas on the facial regions, the proposed approach actually removes "noise" information caused by other non-significant areas of the facial region. It may also remove part of the variation information caused by changes in facial expression, head rotation and illumination. By concentrating on these significant areas, it allows us to extract the most significant facial features from them to represent human faces. These features may improve the classification rate of face recognition systems.

The first step in the proposed technique is to locate facial regions in the face images. The facial feature extraction technique is performed on these facial regions. After feature extraction, the features are selected, fused and classified. Through selection, the significant areas are located, and through classification, the input face image is recognised or verified.

The block diagram of the proposed technique to conduct experiments using separate and combined features on FERET benchmark dataset is depicted in Fig. 2.1. The details are described in the following subsections.

2.2. Locate facial regions

We first locate facial regions on each face image and then we extract features. The experimental face images are extracted from the FERET database. The center coordinate information provided

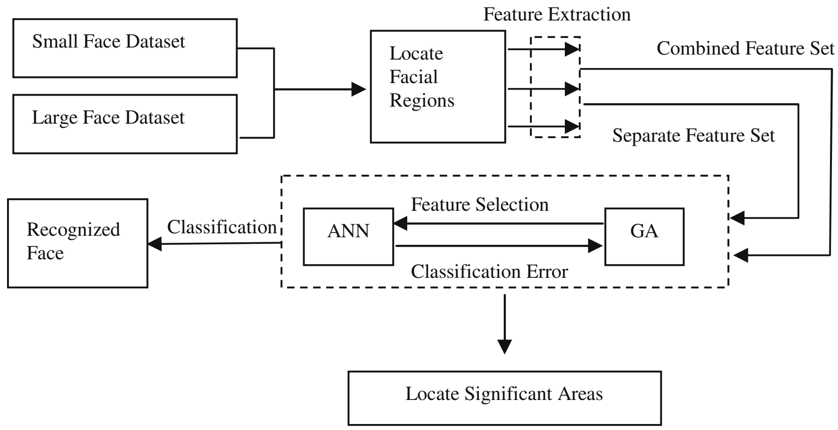

Fig. 2.1. Block diagram of the proposed technique.

for each facial region, such as eye center coordinate, nose tip coordinate and mouth center coordinate, is used and the distance threshold method is applied to locate the local facial regions.

The distance threshold method defines distance thresholds in vertical and horizontal directions for the local facial region. These thresholds decide the size of the facial region. With center coordinate information, the facial region is easy to locate. Based on the images in the experimental database, the distance thresholds are set as follows. The vertical distance threshold is set to 16 and the horizontal distance threshold is set to 30 for the eyes and nose regions. They are set to 12 and 60 separately for the mouth region.

2.3. Average grey level value feature

After locating the facial regions, each facial region was equally divided into small-size rectangle areas. The average grey level value features are extracted from these small rectangle areas. The average grey level value feature can be expressed as

$$g_i = \frac{\sum p(x,y)}{w \times h \times v} \tag{1}$$

where g_i is the average grey level value feature for the small rectangle area i, $p(x,y)$ is grey level value of pixel p inside the rectangle area i, w is the width of the small rectangle area and h is the height. v is the maximum grey level value for the image, here 255 for the experimental database.

After division, the average grey level value features are extracted on these small rectangle areas from left to right, top to bottom. In the experiments, The size of the small rectangle area was chose to be 6×4 ($w = 6$, $h = 4$). Then for the left eye region (same as right eye region and nose region), the size of the extracted feature set becomes 20. For the mouth region, the size of the extracted feature set increases to 30 due to the larger size of the mouth region.

2.4. PCA feature

PCA projection method for face recognition, which is also called eigenface method, is a classical method for face recognition. The simple idea behind eigenface method is to capture the largest variances among a set of face images, and then use this information to encode and compare face images. The advantage of eigenface method is reduction of dimensionality, while maximizing the scatter of all the projected samples. Let $\{X_1, X_2, \ldots, X_N\}$ be as set of N sample images. It takes values in an n-dimensional image space and each image belongs to one of the c classes $\{x_1, x_2, \ldots, x_c\}$. A linear transformation needs to be found to map the original n-dimensional image space into an m-dimensional feature space, where $m < n$. The new feature vector $y_k \in R^m$ is defined by the following equation:

$$y_k = W^{\mathrm{T}} X_k, \quad k = 1, 2, \ldots, N \tag{2}$$

where $W \in R^{n \times m}$ is a matrix with orthonormal columns. W is chosen to maximize the determinant of the total scatter matrix S of the projected samples.

$$S = \sum_{k=1}^{N} (X_k - \mu)(X_k - \mu)^{\mathrm{T}} \tag{3}$$

$$W_{\mathrm{opt}} = \mathrm{argmax}_w |W^{\mathrm{T}} S W| = (w_1, w_2, \ldots, w_m) \tag{4}$$

where N is the number of sample images and μ is the mean image of all the samples. $\{w_i | i=1,2,\ldots,m\}$ is the set of n-dimensional eigenvectors of S corresponding to the m largest eigenvalues. In the experiments, the PCA projection method is applied to local facial regions instead of the whole face images to extract features.

2.5. GA–ANN technique

The GA and ANN-based technique is used to identify the significant areas in each facial region and perform fusion and selection of features for face recognition. In this research, GAs are used to find potential significant features which will generate a higher recognition rate. The areas that contain these significant features are considered to be the significant areas. The chromosomes represent the possible selection of the significant features. Binary encoding is used for the chromosomes, where 1 represents that the feature is selected and 0 represents that the feature is not selected. In one generation, each chromosome is multiplied by the input feature set to generate the input feature vector to ANN. The input feature vector F can be represented as

$$F = CP \tag{5}$$

$$C = (c_1, c_2, \ldots, c_l), c_i \in \{0, 1\} \tag{6}$$

$$P = L + R + N + M \tag{7}$$

where C is a single chromosome, c_i is one gene in the chromosome. l is the length of the chromosome, which is the same as the size of the input feature set P. When testing on the separate feature set from each facial region, P represents the separate feature set. As mentioned in the last section, size of the left eye feature set L is 20, size of the right eye feature set R is 20, size of the nose feature set N is 20 and size of the mouth feature set M is 30. When combining them together, the size of P is 90. Eq. (7) shows the combining feature set P.

The input feature vector F is fed to ANN for classification. An ANN with a single hidden layer is used in this technique. A resilient backpropagation algorithm is used to train the network. The testing classification error is used to calculate the fitness of the corresponding individual in GA. In the reproduction, the 'fittest' individual that achieves the best testing classification rate is retained in the next generation. The chromosomes in each generation of GAs that achieve the best classification rate are recorded. The chromosomes indicate which feature is selected and which is not. After all generations, the total number of times that each feature has been selected for the best classification rate is calculated. All the features are ranked according to the number of times they have been selected. The areas that contain the feature inside top n ranking are the top n significant areas. For the experiments, n is defined as 3.

3. Databases

Three experimental databases were used in this research. All of them are extracted from the FERET benchmark database. The preliminary experimental database is a small subset of FERET database and consists of 13 classes (one class represents one distinct person). In each class, there are four face images. Three of them are randomly chosen for training and the remainder is chosen for testing. The total number of face images in the database is 52. The images are selected carefully in order to have minimum pose variation. Fig. 4.1 shows the example images from the preliminary database. Top 3 rows show training images and bottom row shows testing images.

The advance databases consist of 50 classes, each class represents one distinct person. In the original dataset (DB1) from our previous study, there are four face images in every class. Three of them are randomly selected for training and one for testing. The extended dataset (DB2) includes all the images from DB1 and more images. In DB2, each class has 4–12 images for training and one for testing. There are totally 376 images for training and 50 images for testing in DB2.

Fig. 4.1. Example images of the preliminary database. The top 3 rows show training images and the bottom row shows testing images.

4. Experimental results

This section describes the experimental results on small and large databases. The goal of the experiments is to evaluate the proposed technique and make comparison with the other existing techniques. All experimental databases are extracted from the FERET benchmark database. Section 4.1 presents the experiments which are based on the preliminary database. Section 4.2 describes the advance experiments which are based on larger databases.

4.1. Preliminary results

The preliminary experiments were conducted using preliminary database as described in Section 3. Fig. 4.1 shows the example images from the preliminary database. Top 3 rows show training images and bottom row shows testing images.

At first, the experiments on location of the significant areas using GA–ANN were conducted on each facial region separately. The features used in the experiments were average grey level value features. The extracted average grey level value features from each facial region formed the input feature vector for that region. The size of the small rectangular area for feature extraction was chosen to be 6×4. The size of the extracted feature set L from left eye region was 20. The same size was for the extracted feature set R from right eye region and the extracted feature set N from nose region. For mouth region, the size of the extracted feature set M was 30. L, R, N, M can be expressed in the following equations

$$L = (l_1, l_2, l_3, \ldots, l_{20}), \quad l_i \in (0, 1) \tag{8}$$

$$R = (r_1, r_2, r_3, \ldots, r_{20}), \quad r_i \in (0, 1) \tag{9}$$

$$N = (n_{1,} n_2, n_3, \ldots, n_{20}), \quad n_i \in (0, 1) \tag{10}$$

$$M = (m_1, m_2, m_3, \ldots, m_{30}), \quad am_i \in (0, 1) \tag{11}$$

These extracted feature sets were then fed to GA–ANN separately for selection and classification. To make the experiments consistent, the parameters of GA–ANN were set exactly same for every set of experiments. The generation number was set to 50 and the population number was set to 15. The crossover rate was set to 0.9 and the mutation rate was set to 0.2. The hidden units of ANN were increased from 6 to 44 (each time increased 2 hidden units), and the selections that generated the best recognition rate were recorded. The epoch for ANN was set to 3000.

The best classification results for each facial region feature set are shown in Tables 4.1–4.4. The shadowed cells indicate the highest testing classification rate.

The results given in Tables 4.1–4.4 show that the eye region and the mouth region achieve better recognition rate than the nose region. For the nose region, the best recognition rate is just 76.92% when the hidden units are 34. For left eye region, the best recognition rate is 92.31% when hidden units are 14, 24 and 34. For the right eye region, the best recognition rate is 92.31% when the hidden units are 30 and 44. For mouth region, the best recognition rate is also 92.31% when hidden units are 10 and 36.

When achieving the best recognition rate for each facial region, the corresponding feature selection and combinations that were obtained are shown in Table 4.5. The results given in Table 4.5, the left eye region has four different feature combinations, which contain the same feature $l_{1,}$ $l_{8,}$ $l_{9,}$ l_{20}. The right eye region has two different feature combinations, which contain the same feature $r_{2,}$ $r_{3,}$ $r_{4,}$ $r_{5,}$ $r_{7,}$ $r_{9,}$ $r_{10,}$ $r_{16,}$ r_{19}. The mouth region also has two different

Table 4.1
Best classification results for the left eye feature set

Hidden Units	Feature Selection (GA Chromosomes)	Classification Rate		RMS Error
		Training Rate [%]	Testing Rate [%]	
14	1 0 0 0 0 0 0 1 1 0 0 0 0 1 0 1 0 1 0 1	100	92.31	0.04719
	1 0 0 0 0 0 0 1 1 0 0 0 0 1 0 0 0 1 0 1	97.44	92.31	0.06064
16	1 1 1 0 1 1 1 1 1 1 1 1 0 1 1 1 1 1 1 1	100	84.62	0.03070
	1 1 0 0 0 1 1 1 1 0 0 1 1 1 1 1 0 0 0 1	100	84.62	0.03960
24	1 0 0 0 1 1 0 1 1 0 0 1 1 0 1 1 0 0 1 1	100	92.31	0.03910
30	1 1 1 1 1 0 0 0 0 1 0 0 1 1 1 0 1 1 0 1	100	84.62	0.03758
	1 1 1 1 1 0 0 0 0 1 0 0 1 0 0 0 1 0 0 1	100	84.62	0.03948
34	1 0 0 0 0 1 1 1 1 0 0 0 1 0 0 1 1 1 1 1	100	92.31	0.03698

Table 4.2
Best classification results for the right eye feature set

Hidden Units	Feature Selection (GA Chromosomes)	Classification Rate		RMS Error
		Training Rate [%]	Testing Rate [%]	
14	0 0 0 1 1 0 1 0 0 1 0 0 1 0 1 0 0 1 0 1	100	84.62	0.04686
16	0 0 0 1 1 0 1 1 0 1 1 0 0 0 0 0 0 1 1 0	100	84.62	0.05209
24	1 0 0 1 0 1 1 1 0 1 0 0 1 0 0 1 0 1 1 1	100	84.62	0.04268
	1 0 0 1 0 1 1 1 0 1 0 1 1 0 0 1 0 1 1 1	100	84.62	0.03669
30	0 1 1 1 1 1 1 0 1 1 1 0 0 0 1 1 0 0 1 1	100	92.31	0.03492
36	0 1 0 1 1 0 0 1 0 1 0 1 0 0 0 1 0 1 0 1	100	84.62	0.04036

Table 4.3
Best classification results for the nose feature set

Hidden Units	Feature Selection (GA Chromosomes)	Classification Rate		RMS Error
		Training Rate [%]	Testing Rate [%]	
14	1 0 0 0 0 1 0 0 0 0 0 0 0 1 0 0 1 0 1 1 0	76.92	69.23	0.10628
16	1 1 1 1 1 0 0 0 0 1 1 0 1 0 0 0 0 0 0 1 1	92.31	61.54	0.08497
	1 1 1 1 0 1 0 1 0 0 0 0 1 1 0 0 1 1 1 1 1	87.18	61.54	0.07875
	1 1 1 0 1 1 0 0 1 1 1 0 1 1 0 0 1 1 1 1 1	97.44	61.54	0.06720
22	0 1 0 0 0 1 0 1 0 0 1 0 1 0 1 0 1 0 0 0	87.18	61.54	0.09534
	0 1 0 0 0 1 0 1 0 0 1 0 1 0 1 0 1 0 1 0	94.87	61.54	0.08828
32	1 1 0 0 0 0 0 0 0 1 1 1 0 1 0 0 0 0 0 1 1	89.74	69.23	0.07957
34	1 1 0 1 0 0 0 0 1 1 0 0 1 0 1 1 0 1 1 1	97.44	76.92	0.06574

feature combinations, which contain the same features m_2, m_3, m_{10}, m_{12}, m_{16}, m_{20}, m_{22}, m_{29}. The nose region just has one feature combination.

All feature sets were combined together to feed to GA–ANN again for the experiments. The size of the input feature vector increased to 90. The parameters of GA–ANN were set exactly the same as those of the previous experiments. Table 4.6 lists the best classification results achieved. The recognition rate is improved to 100%. When the recognition rate is 100%, these selected features were added together to locate the most selected features. The

Table 4.4
Best classification results for the mouth feature set

Hidden Units	Feature Selection (GA Chromosomes)	Classification Rate		RMS Error
		Training Rate [%]	Testing Rate [%]	
10	011101000111101100011100100011	94.88	92.31	0.056066
14	011110000100100101110000100100	100	84.62	0.048529
18	100011011111011010011000110111	100	84.62	0.03591
30	111111100111110001111101000111	100	84.62	0.029281
	111111100111111100111111100011	100	84.62	0.028075
	111111100111111100110011101111	100	84.62	0.02902
32	011011000000001001010010101111	100	84.62	0.038419
36	011010010101010101110100011010	100	92.31	0.03661

Table 4.5
Feature selections for achieving the best recognition rate

Facial region	Hidden units	Feature selection
Left eye region	14	$l_1, l_8, l_9, l_{14}, l_{16}, l_{18}, l_{20}$
		$l_1, l_8, l_9, l_{14}, l_{18}, l_{20}$
	24	$l_1, l_5, l_6, l_8, l_9, l_{12}, l_{13}, l_{15}, l_{16}, l_{19}, l_{20}$
	34	$l_1, l_6, l_7, l_8, l_9, l_{13}, l_{16}, l_{17}, l_{18}, l_{19}, l_{20}$
Right eye region	30	$r_2, r_3, r_4, r_5, r_6, r_7, r_9, r_{10}, r_{11}, r_{15}, r_{16}, r_{19}, r_{20}$
	44	$r_2, r_3, r_4, r_5, r_7, r_9, r_{10}, r_{14}, r_{16}, r_{19}$
Nose region	34	$n_1, n_2, n_4, n_9, n_{10}, n_{13}, n_{15}, n_{16}, n_{18}, n_{19}, n_{20}$
Mouth region	10	$m_2, m_3, m_4, m_6, m_{10}, m_{11}, m_{12}, m_{13}, m_{15}, m_{16}, m_{20}, m_{21}, m_{22}, m_{25}, m_{29}, m_{30}$
	36	$m_2, m_3, m_5, m_8, m_{10}, m_{12}, m_{14}, m_{16}, m_{18}, m_{19}, m_{20}, m_{22}, m_{26}, m_{27}, m_{29}$

Table 4.6
Classification results for combined feature set

Hidden units	Training classification rate (%)	Testing classification rate (%)	RMS error
8	100	100	0.030384
24	100	100	0.008929
38	100	100	0.009836
44	100	100	0.01825

top 10 selected features are shown in Table 4.7. Most of these features are concentrated in the eye region and there is no feature coming from the nose region.

Table 4.7
Top 10 most selected features

Rank	Features
1	l_{20}
2	l_{11}, r_5
3	m_{21}
4	r_1
5	m_{28}
6	r_3, r_4
7	r_{19}
8	m_{27}
9	l_{19}
10	r_6

4.2. Advance experiments

The preliminary experiments achieved very good results. This indicates that the proposed technique is promising. Since the preliminary database is relatively small, the proposed technique needs to be investigated on much larger databases. Another two databases are set up to conduct the experiments, referred to as Databases 1 and 2. Same as the preliminary database, both Databases 1 and 2 are extracted from the FERET database and consist of 50 classes each. In Database 1, there are 150 face images for training and 50 face images for testing. Database 2 includes all the images from Database 1 and increases the training set. In Database 2, there are totally 376 face images for training and 50 face images for testing. Section 4.2.1 presents the experimental results from Database 1 and Section 4.2.2 presents the results from Database 2.

4.2.1. Database 1 results

There are four face images per class in Database 1. Three of them are randomly selected for training and the left one is selected for testing. Example images from Database 1 could be found in Fig. 4.2. Two different sets of experiments were conducted on Database 1. In the first set of experiments, the average grey level value features were investigated. In the second set of experiments, the PCA features were investigated. Section 4.2.1.1 describes the experiments using average grey level value features. The experiments using PCA features are explained in Section 4.2.1.2.

4.2.1.1. Average grey level value features. For the experiments using average grey level value features, two different sizes of small rectangular areas for feature extraction were investigated. In the experiments, the size of the small rectangular area was firstly set to 6×4 and then set to 10×4. Section Section 4.2.1.1.1 presents the results when the size of the small rectangular area is 6×4.

4.2.1.2. Small rectangular area size is 6×4. When the size of the small rectangular area for feature extraction is 6×4, the GA–ANN technique was firstly tested on each facial region feature set separately. During the experiments, the hidden units of ANN were increased from 8 to 64 (an increment of 4 hidden units each time), and the selections that generated the best recognition rate were recorded. To make the experiments consistent, the other parameters of GA–ANN were set exactly same for every experiment. The generation number was set to 40 and the population

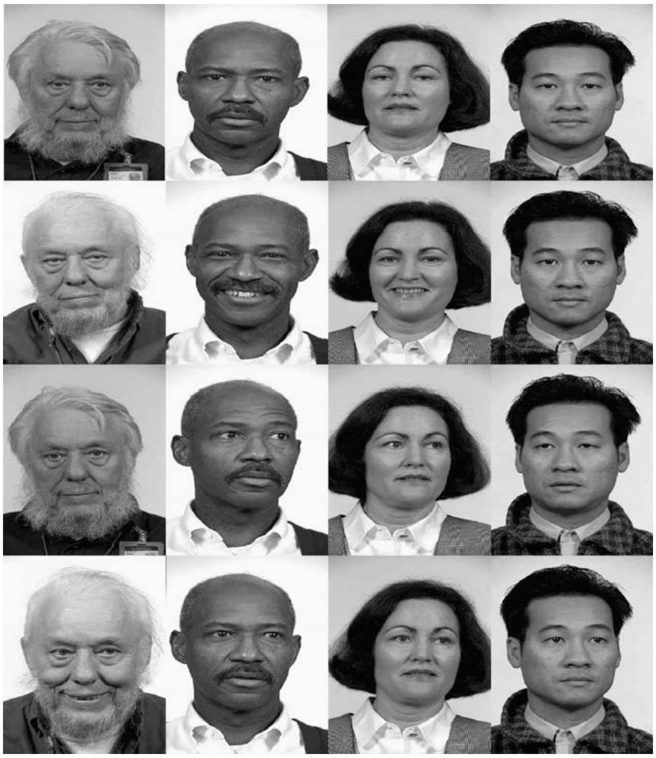

Fig. 4.2. Example images from Database 1. Top three rows show training images and bottom row shows testing images.

number was set to 10. The crossover rate was set to 0.9 and the mutation rate was set to 0.2. The epoch for ANN was set to 10,000. Tables 4.8–4.11 list the best classification results achieved for each facial region feature set.

In the above tables, the shadowed cells indicate the highest testing classification rate. The highest testing classification rate for the left eye feature set is 54% (hidden units 44, 56, 60), for the right eye feature set is 62% (hidden units 28), for the nose feature set is 38% (hidden units 52) and for the mouth feature set is 70% (hidden units 36). The results given in Tables 4.8–4.11 show that the mouth region alone achieved the best classification rate, while the nose region achieved the worst classification rate.

The extracted average grey level value features from each facial region were combined together to form the input feature vector for GA–ANN. The size of the input feature vector increased to 90. The other parameters of GA–ANN were set exactly the same as those of the experiments using the separate facial feature set. The feature combination sequence was left eye, right eye, nose and mouth.

Table 4.12 lists the results of the combined feature set that achieved above 80% testing classification rate on Database1. The results given in the table show that the best testing classification rate is 86% (hidden units 40, 56), and the best training classification rate is 100%. The shadowed cells indicate the best testing classification rate. The combined feature set outperformed the separate feature set from each facial region and improved the classification rate significantly.

To investigate the effects of feature combination sequence on the recognition rate, the feature combination sequence was reversed to mouth, nose, left eye and right eye to form a new input vector. Then the same experiments under the same parameters were conducted. Table 4.13 lists the best classification results of the combined feature set in the reverse order. The best recognition rate is still 86% (hidden units 64).

When the recognition rate is 86% (hidden units 40, 56) for the combined feature set in the original order, the total selection times of each feature were calculated. By mapping the total selection times of each feature to its corresponding extraction area, Fig. 4.3 is generated. In Fig. 4.3, the shaded areas are the areas that contain the top selected features. These areas are considered to be the significant areas. There are totally 36 areas. Among these areas, there are 9 areas from the left eye region, 7 areas from right eye region, 5 areas from the nose region and 15 areas from the mouth region.

The results in Table 4.13 show that the best recognition rate is still 86% when the feature combination sequence is reversed. When the recognition rate is 86% (hidden units 64), the total selection times of each feature were calculated. Similarly, by mapping the total selection times of each feature to its corresponding extraction area, Fig. 4.4 is generated. In Fig. 4.4, the shaded areas are the areas that contain the top selected features. There are total 49 areas. Among these areas, there are 7 areas from left eye region, 12 areas from right eye region, 11 areas from nose region and 19 areas from mouth region.

4.2.1.3. PCA features. PCA features are extracted separately from each facial region, and then combined together to form the input feature vector for GA–ANN. After feature extraction, the sequence of feature combination is left eye, right eye, nose and mouth.

Because we do not know how many eigenvectors should be suitable for encoding the face images, a different number of

Table 4.8
Best classification results for the left eye feature set

Hidden Units	Feature Selection (GA Chromosomes)	Classification Rate		RMS Error
		Training Rate [%]	Testing Rate [%]	
12	1 0 0 0 1 1 1 1 1 1 1 1 0 0 0 1 1 1 0 0	82	48	0.090757
32	1 1 0 1 1 0 0 1 1 0 0 1 1 1 0 1 1 0 0 0	100	52	0.065463
40	1 1 1 1 1 0 0 0 1 0 1 0 0 1 1 1 0 1 1 0	100	48	0.055730
44	1 1 1 1 1 0 0 0 0 1 0 0 1 1 1 1 1 1 0 1	100	54	0.055398
48	1 1 1 1 1 0 0 0 0 1 0 0 1 1 1 0 1 1 0 1	100	50	0.055502
56	1 1 1 1 1 1 1 1 0 1 1 0 1 1 1 0 1 1 1 0	100	54	0.04531
60	1 1 1 1 1 0 1 1 1 0 1 1 1 1 1 0 1 1 0 1	100	54	0.045535
64	1 1 1 1 1 0 0 0 0 1 0 0 1 1 1 0 1 1 0 1	100	52	0.050977

Table 4.9
Best classification results for the right eye feature set

Hidden Units	Feature Selection (GA Chromosomes)	Classification Rate		RMS Error
		Training Rate [%]	Testing Rate[%]	
16	1 1 1 1 1 0 0 1 0 1 0 0 1 1 1 0 1 1 0 1	100	56	0.075269
24	1 1 1 1 1 0 0 0 0 1 1 1 0 0 0 1 0 1 0 1	98	56	0.067813
28	1 1 1 1 1 0 0 0 1 1 0 0 0 1 0 1 0 0 0 0	98.67	62	0.067738
36	1 1 1 1 1 0 0 0 0 1 0 0 1 0 0 1 0 0 1 0	99.33	60	0.061682
44	1 1 1 1 1 0 0 0 0 1 0 0 1 1 0 1 0 1 0 1	100	56	0.052249
52	1 1 0 1 1 0 1 1 0 1 0 0 0 0 0 1 0 0 0 1	99.33	60	0.058238
60	1 1 1 1 1 0 0 0 0 1 0 0 0 0 0 1 0 0 0 1	99.33	58	0.062595
64	1 1 1 1 1 0 0 1 0 1 0 0 0 0 0 1 0 0 0 1	99.33	58	0.057562

Table 4.10
Best classification results for the nose feature set

Hidden Units	Feature Selection (GA Chromosomes)	Classification Rate		RMS Error
		Training Rate [%]	TestingRate[%]	
20	1 0 1 0 0 0 1 1 1 1 1 0 1 1 0 1 1 1 0 1	82	34	0.083747
32	1 1 1 1 1 1 1 0 1 0 0 1 0 1 0 1 0 1 1 1	94.67	32	0.074111
36	1 1 1 0 1 0 1 0 1 1 1 0 0 0 1 0 1 1 1 0	91.33	32	0.078527
44	1 0 0 1 0 1 1 0 0 0 0 1 0 0 1 0 0 1 0 1	87.33	34	0.087115
48	1 1 0 1 0 0 0 0 0 0 0 0 1 1 0 0 1 0 1 1	84.67	30	0.084655
52	1 1 1 1 1 1 1 1 0 1 1 0 1 1 1 0 1 1 0 1	98.67	38	0.06674
60	1 0 0 1 0 1 1 0 0 0 1 0 0 0 1 0 1 1 0 1	88	34	0.081849
64	1 0 1 1 1 0 0 0 0 1 0 0 1 1 1 0 1 1 0 1	97.33	32	0.075388

Table 4.11
Best classification results for the mouth feature set

Hidden Units	Feature Selection (GA Chromosomes)	Classification Rate		RMS Error
		Training Rate [%]	Testing Rate [%]	
32	101101110100011010100100100110	100	62	0.055177
36	111111111110111001011000110010	100	70	0.047729
40	111111001110011010000100100010	100	66	0.048634
44	001110110101100101100100101010	100	62	0.047705
48	001110110100011010100100011010	100	66	0.048532
52	111111110100011010100100100101	100	66	0.042634
	111111110010011010100100111010	100	66	0.04244
56	110001001101101110011000110010	100	68	0.042558
60	001101001111101010101001010110	100	62	0.04191
	001101001111101001010101010110	100	62	0.041762
	001101010011101001010101010110	100	62	0.043019
	001010101111101001010101010110	100	62	0.0413
64	101001110000010101101000111101	100	64	0.042305

Table 4.12
Best classification results for combined feature set in original order

Hidden Units	Feature Selection (GA Chromosomes)	Classification Rate		RMS Error
		Training Rate [%]	Testing Rate [%]	
36	11010000010010000011110011101101110001011011100100001011011010000111101111110110110111 0111	100	80	0.03366
	1101000001001000001111001110110111 00010110111001011101001001011110000010 00000100100100001111	100	80	0.03500
40	00101111101001111100001111001111101 00100001010001100111110011011010011101 1100111111111101101	100	86	0.02964
48	1000101100001101001111001110110111 00010110111001011101001001011110000010 00000100100100001000	100	82	0.02819
56	1010111111011011101111001110110111 11101001000110100010100101011111100110 0111110111111101010	100	86	0.01939
60	0011110000110001110110100110111111 11000011010001011000110010010110101 00 0001001010011011111	100	84	0.02054

Table 4.13
Best classification results for combined feature set in the reverse order

Hidden Units	Feature Selection (GA Chromosomes)	Classification Rate		RMS Error
		Training Rate[%]	Testing Rate[%]	
28	1010010111111111010001101100111010 1 10101001000000100110110000100001111011 1 0011111111111101101	100	80	0.038631
40	0011101101000110101001000010110110 0 01111101000000000100100010000110110100 0 00101110000011100	100	80	0.033954
52	0001101011000110101001001001010101 1 00011110000010110001100100110110001001 01101111000001101	100	82	0.023553
56	100111010110100101010111000010100 11 10100111010011110110100100001101111110 00001000000010101	100	82	0.021568
60	0011110101101010101010001111010110 0 01011100101100110110111011110010010111 11011110111100001	100	80	0.017561
64	1101110101101010001111001110110111 0 00110011010011110111011011001101101000 00100110000011110	100	86	0.016062
	1101110101101010101111001110110111 0 00110011010011110110111011001101101000 00100110000011110	100	86	0.015144
68	0110001010010101010101110000101001 1 10100100101100001001000100001101101000 00100001111100001	100	82	0.01995
	1001110101101010101010001111010110 0 01011011010011110110111011110010010111 11011110000011110	100	82	0.015121

eigenvectors was evaluated in the experiments. The experiments using 10 eigenvectors, 14 eigenvectors, 18 eigenvectors and 22 eigenvectors were conducted. The parameters of GA–ANN were set exactly same for every experiment. The generation number was set to 40 and the population number was set to 10. The cross-over rate was set to 0.9 and the mutation rate was set to 0.2. The epoch for ANN was set to 10000. The hidden units were increased from 8 to 68 (an increment of 4 hidden units each time).

Table 4.14
Best classification results for 10 eigenvectors

Hidden Units	Feature Selection (GA Chromosomes)	Classification Rate		RMS Error
		Training Rate [%]	Testing Rate [%]	
12	000000101110010011000010110111101 0111001	100	60	0.080433
	111111010100101000111100111000110 0110011	98.67	60	0.076291
36	000000101110010010110011010010110 0110011	100	64	0.043994
52	100001100111111110000010000010110 00110010	100	62	0.035194
	100001100111111111100100000010110 00110010	100	62	0.033143
56	000000101110010010110011010010110 0110011	100	66	0.033341

Table 4.15
Best classification results for 14 eigenvectors

Hidden Units	Feature Selection (GA Chromosomes)	Classification Rate		RMS Error
		Training Rate [%]	Testing Rate [%]	
20	0000110010110011010010110011001110000110011111111000000100	100	68	0.059964
	1111011110110011010010110011001110000110011111111000011011	100	68	0.055764
40	0000110010110011010010110011001110000110011111111001010 01	100	74	0.038337
48	0000110010110011010010110011011110000110011111111000000100	100	76	0.034414
60	0000110010110011010010101111011110000110011111111011000 01	100	78	0.026711
64	0000110010110011010101011100010011100 001100111111100010010	100	78	0.027607

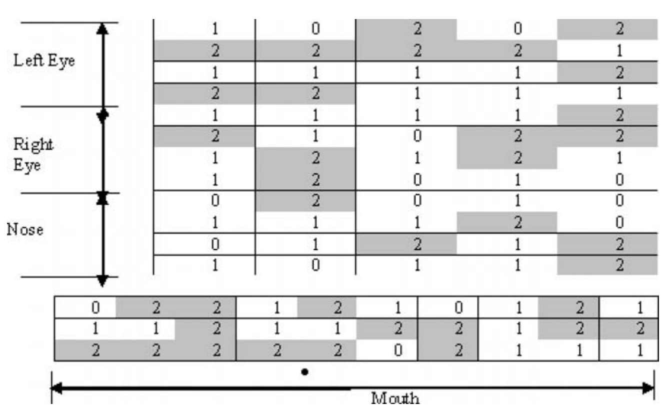

Fig. 4.3. Significant areas in facial regions (original order combination).

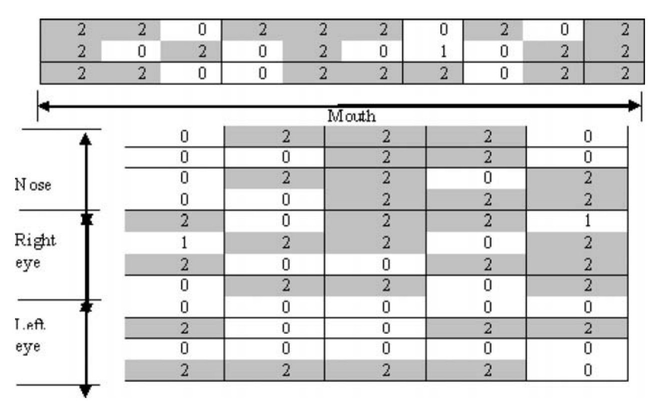

Fig. 4.4. Significant areas in facial regions (reverse order combination).

Tables 4.14–4.16 present the best classification results for a different number of eigenvectors. The results given in Table 4.14 show that the best testing classification rate for 10 eigenvectors is 66% when the hidden units are 56. The corresponding best training classification rate is 100%.

Table 4.15 shows that the best testing classification rate for 14 eigenvectors is 78% when the hidden units are 60 and 64. The corresponding best training classification rate is 100%.

Table 4.16 shows that the best testing classification rate for 18 eigenvectors is 80% when the hidden units are 60. The corresponding best training classification rate is 100%.

4.2.2. Database 2 results

In Database 2, each class has 4–12 images for training and one for testing. Because the combined feature set (using average grey level value features) achieved much better results on Database 1,

Table 4.16
Best classification results for 18 eigenvectors

Hidden Units	Feature Selection (GA Chromosomes)	Classification Rate		RMS Error
		Training Rate [%]	Testing Rate [%]	
16	11011110101100101010101111110010000 1010001000110101000110110100000000000	100	70	0.064818
32	11111001101111101010100100111101111 0101110110110101000110010011101010100	100	72	0.040375
40	00001000111010101010111010111010000 0110100000101100000111100111011011100	100	74	0.036662
48	11101100011010101010100101001010011 0110100000101100000111100111011011100	100	68	0.030918
56	10010111100000111101011010111010000 0110100000101100000111100111011011100	100	74	0.026269
60	10101011101010101001001010111110000 0110100000110011011100111111111101101	100	80	0.024114
64	00010011100101010101011010111010000 0110100000101100000110110011010100001	100	74	0.027777

Table 4.17
Best classification results (Database 2) of hidden units 40 feature selection (Database 1)

Hidden Units	Training Classification Rate	Testing Classification Rate
30	100%	88%
42	100%	94%
48	100%	88%
52	100%	90%
54	100%	92%
60	100%	90%
66	100%	88%
74	100%	94%
76	100%	90%
82	100%	88%
86	100%	94%

Table 4.18
Best classification results (Database 2) of hidden units 56 feature selection (Database 1)

Hidden Units	Training Classification Rate	Testing Classification Rate
26	100%	88%
38	100%	92%
44	100%	92%
54	100%	90%
58	100%	94%
68	100%	90%
72	100%	92%
78	100%	88%
80	100%	88%
88	100%	90%

so only the combined feature set (using average grey level value features) experiments were conducted on Database 2. To make the experiments faster, the best feature selection from the previous experiments on Database 1 was used directly to train and test the ANN. The epoch was increased to 15,000 because there are more face images in the database. More hidden units were also used in the experiments.

When the size of the small rectangular area was 6×4, the hidden units 40 and 56 feature selections achieved the best recognition rate on Database 1. These two feature selections were directly used in the experiments on Database 2. The results based on the hidden units 40 feature selection are listed in Table 4.17. The results based on the hidden units 56 feature selection are presented in Table 4.18. The results given in both tables, the highest recognition rate is improved to 94%.

When the size of small rectangular area was 10×4, the hidden units 44 feature selection achieved the best recognition rate on Database 1. Table 4.19 shows the best classification results using hidden units 44 feature selection on Database 2. The results given

Table 4.19
Best classification results of hidden units 44 feature selection

Hidden Units	Training Classification Rate	Testing Classification Rate
28	99.20%	90%
30	99.73%	88%
34	100%	88%
36	99.47%	90%
38	100%	94%
42	100%	90%
50	99.73%	90%
52	100%	92%
54	100%	94%
56	100%	94%
62	100%	90%
64	100%	92%
66	100%	92%

in Table 4.19 show that the highest recognition rate is also improved to 94%.

5. Comparative analysis

The results obtained in this research are compared to the results of the other methods mentioned in a recent study (Zhang et al., 2004). The authors (Zhang et al., 2004) also extracted a dataset from the FERET database as their experimental database, which has 927 images corresponding to 119 persons. Three different methods are experimented on this dataset. These methods include kernel associative memory (kAM) method which is proposed in their study, PCA-nearest-neighbor method and a simple NN-based template matching method termed ARENA. In this study, we compared with the highest classification rate achieved in their (Zhang et al., 2004) study. They conducted two sets of experiments similar as in our research: the first one used 3 images per class for training and the second one used 4 images per class for training.

Fig. 5.1 shows the comparison of the best recognition rates between our DB1 (Database 1) experimental results and their first set results. Fig. 5.2 shows the comparison of the best recognition rates between our DB2 (Database 2) experimental results and their second set results. Both figures show that our approach achieves a better recognition rate.

6. Conclusions

We have presented a feature selection and fusion technique for face recognition in this paper. The GA for feature selection and ANN for feature classification are incorporated in the proposed technique. The technique performs fusion and selection of facial features for face recognition. The significant areas inside each facial region are located through the feature selection.

The FERET benchmark database is adopted to evaluate and compare the proposed technique with the existing techniques. Three different databases were used in the experimental investigation and all of them were extracted from the FERET benchmark database. Database 2 is the largest database, containing 50 classes and 426 face images. The experiments are conducted on Cluster machine at Central Queensland University.

The preliminary experiments were conducted simply to pre-test the proposed technique. The experiments investigated the separate facial region feature set and the combined feature set using the average grey level value features. The preliminary results were promising. The left eye feature set, right eye feature set and mouth feature set all achieved the highest recognition rate of 92.31%. The nose feature set just achieved the highest recognition rate 76.92% and was the worst performer. The combined feature set outperformed the separate facial region feature set by improving the recognition rate to 100%. On Database 1, many experiments were conducted to perform further investigation. The different size of the feature extraction area, the different feature extraction technique and the different sequence for feature combination were considered in the experimental investigation. When average grey level value features were used and the size of the small rectangular area was 6×4, the left eye feature set achieved 54% recognition rate, right eye feature set achieved 62% recognition rate, nose feature set achieved 38% recognition rate and mouth feature set achieved 70% recognition rate (see Table 5.1). The mouth feature set was the best performer and the nose feature set was the worst performer. The combined feature set outperformed the separate facial region feature set by achieving 86% recognition rate. The combination sequence did not affect the recognition rate for combined feature set. For significant areas, the mouth region contributed the

Table 5.2
Combined feature set results on DB1 (Database 1)

Hidden Units	Training Rate [%]	Testing Rate [%]
40	100	86
44	100	82
48	100	82
56	100	86
60	100	84
64	100	84
68	100	82

Table 5.3
Combined feature set results on DB2 (Database 2)

Hidden Units	Training Rate [%]	Testing Rate [%]
26	99.73	88
30	100	88
38	100	88
42	100	94
50	100	90

Table 5.1
Separtate facial region feature set results on DB1

	Hidden Units	Training Rate [%]	Testing Rate [%]
Left Eye Region	44	100	54
	56	100	54
	60	100	54
	64	100	52
Right Eye Region	20	97.33	60
	28	98.67	62
	36	99.33	60
	52	99.33	60
Nose Region	20	82	34
	44	87.33	34
	52	98.67	38
	56	96.67	36
Mouth Region	36	100	70
	40	100	66
	48	100	66
	56	100	68

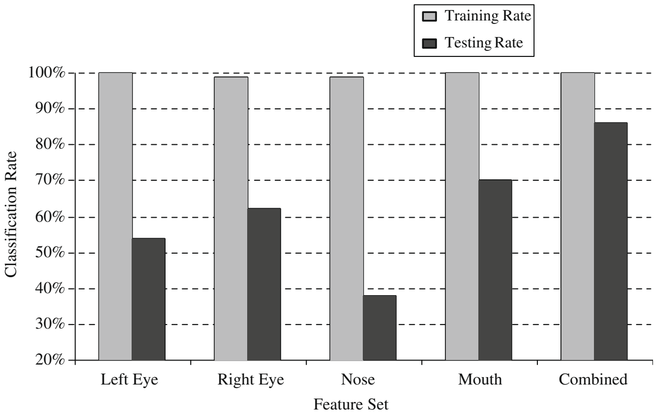

Fig. 5.1. Classification rate comparison between different feature sets.

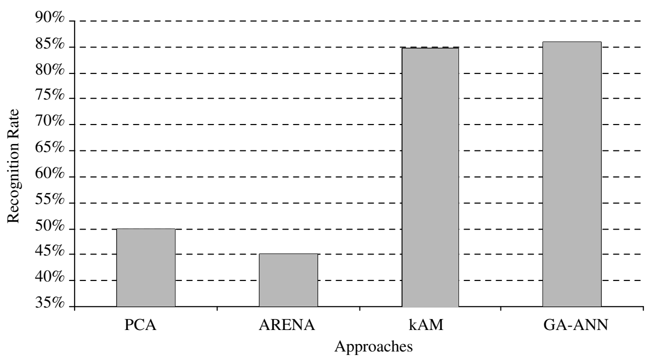

Fig. 5.2. Comparison with other approaches. (3 images per class for training set).

most and the nose region contributed the least. When the size of the small rectangular area was increased to 10×4, the left eye feature set achieved 54% recognition rate, the right eye feature set achieved 56% recognition rate, the nose feature set achieved 36% recognition rate and the mouth feature set achieved 64%. The mouth feature set was still the best performer and the nose feature set was still the worst performer. The combined feature set improved the recognition rate to 86% when compared to the separate facial region feature set (see Table 5.2). The combination sequence slightly affected the recognition rate. The original order combination achieved of 84% recognition rate while the reverse order combination achieved slightly higher recognition rate 86%. For significant areas, the mouth region still contributed the most and the nose region contributed the least. The above results indicate the mouth region is the most important facial region. Combination of facial features from each facial region is much more useful in improving recognition rate compared to just one facial region feature set. A different number of eigenvectors was used for PCA features experiments. The 18 eigenvectors achieved the highest recognition rate 80%. The average grey level value features (combined feature set) outperformed the PCA features by 6%.

The experiments on Database 2 were conducted by just using the combined feature set of average grey level value features. The recognition rate was improved to 94% (see Table 5.3). The experimental results of the proposed approach were also compared with the results of the other three approaches based on FERET database: PCA, ARENA and kAM. Fig. 5.2 was based on Database 1 results and the proposed approach improved the recognition rate by 1.3% compared to kAM method, 36% compared to PCA method and 41% compared to ARENA method. Fig. 5.3 was based on Database 2 results and the proposed technique improved the recognition rate by 2.4% compared to kAM method, 43.2% compared to PCA method and 48.3% compared to ARENA method. The proposed technique achieved the highest recognition rate among the existing techniques based on FERET database.

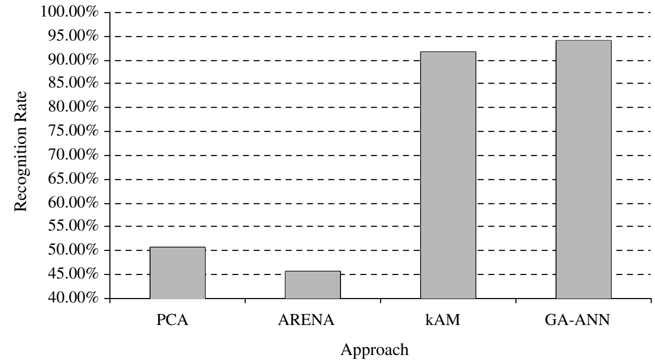

Fig. 5.3. Comparison with other approaches. (4 or more images per class for training set).

References

Bala, J., Huang, J., Vafaie, H., DeJong, K., & Wechsler, H. (1995). Hybrid learning using genetic algorithms and decision trees for pattern classification. *Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence, 1*, 719–724.

Belhumeur, P. N., Hespanha, J. P., & Kriegman, D. J. (1997). Eigenfaces vs. fisherfaces: Recognition using class specific linear projection. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 19*, 711–720.

Brunelli, R., & Poggio, T. (1992). Face recognition through geometrical features. *Proceedings of ECCV, 92*, 792–800.

Chellappa, R., Wilson, C. L., & Sirohey, S. (1995). Human and machine recognition of faces: A survey. *Proceedings of the IEEE, 83*, 705–740.

Fang, Y., Tan, T., & Wang, Y. (2002). Fusion of global and local features for face verification. *IEEE International Conference on Pattern Recognition, 2*, 382–385.

Li, S. Z., & Lu, J. (1999). Face recognition using the nearest feature line method. *IEEE Transactions on Neural Networks, 10*, 439–443.

Liu, Q., Tang, X., Lu, H., & Ma, S. (2004). Kernel scatter-difference based discriminant analysis for face recognition. In *International 17th conference on pattern recognition* (Vol. 2, pp. 419–422).

Liu, C., & Wechsler, H. (2000). Evolutionary pursuit and its application to face recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 22*(6), 570–582.

Liu, C., & Wechsler, H. (2003). Independent component analysis of Gabor features for face recognition. *IEEE Transactions on Neural Networks, 14*, 919–928.

Phillips, P. J., Wechsler, H., Huang, J., & Rauss, P. (1998). The FERET database and evaluation procedure for face recognition algorithms. *Image and Vision Computing, 16*(5), 295–306.

Sun, Y., & Yin, L. (2005). A genetic algorithm based feature selection approach for 3D face recognition. In *Biometric consortium conference*. USA.

Turk, M., & Pentland, A. (1991). Eigenfaces for recognition. *Journal of Cognitive Neuroscience, 3*, 71–86.

Wiskott, L., Fellous, J. M., & Malsburg, C. (1997). Face recognition by elastic bunch graph matching. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 19*, 775–779.

Yan, S., He, X., Hu, Y., Zhang, H., Li, M., & Cheng, Q. (2004). Bayesian shape localization for face recognition using global and local textures. *IEEE Transactions on Circuits and Systems for Video Technology, 1*(14), 102–113.

Zhang, B., Zhang, H., & Ge, S. (2004). Face recognition by applying wavelet subband representation and kernel associative memory. *IEEE Transactions on Neural Networks, 15*, 166–177.