Image Region Selection and Ensemble for Face Recognition

Xin Geng and Zhi-Hua Zhou

National Laboratory for Novel Software Technology, Nanjing University, Nanjing 210093, China

E-mail: {gengx, zhouzh}@lamda.nju.edu.cn

Received MONTH DATE, YEAR. *

Abstract In this paper, a novel framework for face recognition, namely Selective Ensemble of Image Regions (SEIR), is proposed. In this framework, all possible regions in the face image are regarded as a certain kind of features. There are two main steps in SEIR: the first step is to automatically select several regions from all possible candidates; the second step is to construct classifier ensemble from the selected regions. A realization of SEIR based on multiple eigenspaces, namely SEME, is also proposed in this paper. SEME is analyzed and compared with eigenface, PCA + LDA, eigenfeature, and eigenface + eigenfeature through experiments. The experimental results show that SEME achieves the best performance.

Keywords Face recognition; Region Selection; Multiple Eigenspaces; Ensemble learning; Selective Ensemble

1 Introduction

Human faces are complex, changeful and high dimensional patterns. Although it is toilless for human beings to recognize familiar faces, face recognition is a formidable task for machines. Even so, because of the vast potential applications, face recognition has become an active research area of computer vision and pattern recognition for decades.

One of the most representative face recognition techniques is eigenface [1] [2], which is a subspace method based on Principal Component Analysis (PCA) [3]. It tries to find an eigenspace that mostly keeps the variation of face images and then recognize faces in the eigenspace. It has gained so great success that it has become a de facto standard and a common performance benchmark in face recognition [4].

Eigenface is based on the global feature of face images, i.e. it uses the whole face image as training data. When there are gross variations in the input images that greatly affect the global feature, eigenface may fail [5]. However, in this situation, local features such as eyes, mouth and nose are often less affected. So these local features can help recognize faces. Brunelli and Poggio [6] proposed a template-matching method that used both global features and local features. They used four masks respectively to get the regions of eyes, nose, mouth and the whole face. The masks of eyes, nose and mouth are shown in Figure 1 (a). They claimed that local features could achieve better performance than global features. Pentland et al. [5] extended eigenface to local features, getting eigeneyes, eigennoses and eigenmouths, all of which are called eigenfeatures. The templates used to train eigenfeatures are shown in Figure 1 (b), where the white rectangles show the regions of four local features, i.e. left eye, right

*REGULAR PAPER

Supported by the National Science Foundation of China under Grant No. 60325207, 60496320, the Fok Ying Tung Education Foundation under Grant No. 91067, and the Excellent Young Teachers Program of MOE of China.

(a) The masks of facial features [6]

(b) Facial feature templates [5]

Figure 1: Manually extracted facial feature regions

eye, nose and mouth. They indicated that the eigenfeatures alone were sufficient in achieving a high recognition rate equal to that of the eigenface, and the combination of eigenface and eigenfeatures could achieve even better performance. These results illuminate at least two facts: first, the performance of local features is not worse than that of global feature; second, using multiple features may lead to better performance.

Each feature, no matter global or local, is in fact a region in the face image. For the convenience of analysis and computing, the region is often defined as a rectangle in the image. For instance, eigenface uses the whole face image as global feature, and eigenfeature uses the rectangular region around a certain facial feature, such as eye, nose or mouth, as local feature. However, the extension and position of the local feature are empirically determined by operators. Moreover, there are no reliable criterion to determine which facial feature and how many of them should be used. Thus the performance of those local-feature-based methods greatly depends on the experience of the operators.

Suppose all possible regions in the face image are defined as a certain kind of features, then both the global and local features can be regarded as features selected from all possible candidates in the face image. From this point of view, face recognition can be realized in two steps: the first step is to select one or more features from all possible feature candidates; the second step is to train a classifier for face recognition based on the selected features. One problem lying in most existing face recognition techniques, no matter global-feature-based or local-feature-based ones, is that the feature selection procedure is empirically performed by the operators, consciously or unconsciously. To remove such an unstable factor, an automatic image region selection algorithm for face recognition is proposed in this paper. Based on this selection algorithm, a novel framework for face recognition, namely Selective Ensemble of Image Regions (abbreviated as SEIR), and a realization of SEIR, namely Selective Ensemble of Multiple Eigenspaces (abbreviated as SEME), are also proposed.

The rest of this paper is organized as follows. In section 2, the image region selection algorithm is proposed. In section 3, the framework of SEIR and a realization of SEIR, SEME is introduced. In section 4, experiments are reported. The analysis of the properties of SEME, the exploration of a faster variant of the region selection algorithm and the investigation of the generalization of the selected regions are also included in section 4. Finally in section 5, conclusions are drawn and several issues for future work are indicated.

2 Image Region Selection

Using multiple features in face recognition can be re-explained from another point of view. In fact, each feature can be used alone to classify faces, i.e. each feature can be used to train a weak classifier. Then the multi-feature methods can be regarded as special ensemble learning methods. If the definition of features is extended to all possible rectangular regions in the face image, then using several features for face recognition can be regarded as selective ensemble learning [7]. As Zhou et al. [7] indicated, selecting some classifiers to constitute an ensemble may be better than using all of them. The crux of the matter is the selection procedure.

There exist several works on feature selection for face recognition, such as Local Feature Analysis (LFA) [8], and the AdaBoosted Gabor Features [9]. However, until now there is no work on region selection for face recognition. When viewing all possible rectangular regions in the face images as a kind of features, region selection becomes a nature idea to improve the performance of face recognition systems, just like other feature selection procedures [8] [9]. Nevertheless, neither eigenface nor eigenfeature seems to perform any selection procedure. In fact, the selection is unconsciously performed by the operators. It is up to the operators to determine which regions and how many of them should be used. Based on the common sense that facial features, such as eyes, nose and mouth, are crucial in face recognition, most algorithm designers choose to use the image regions around those salient facial features. But this is a very rough selection principal, which has at least two defects. First, although these facial features are salient in human faces, they are not guaranteed to be the most discriminative features. Second, even these features are the most discriminative ones, no one knows how to take full advantage of them. For example, it is uncertain whether the two eyes should be put into the same rectangle like Figure 1 (a), or different rectangles like Figure 1 (b). Consequently, automatic image region selection algorithm should be designed for face recognition.

Since multi-feature methods can be viewed as special ensemble learning process, results from the ensemble learning area may be inspirational for the design of region selection algorithm. According to Krogh and Vedelsby [10], the generalization error of an ensemble consists of two parts, i.e. the average generalization error and the average ambiguity of the component classifiers. Therefore, in order to get a strong ensemble, endeavor should be devoted to the training of accurate and diverse classifiers. Similarly, in order to select suitable regions for face recognition, attention should be focused on those accurate and diverse ones. Here "accurate" and "diverse" actually refer to the weak classifiers trained from the regions. But in order to emphasize that the selection algorithm is independent of the classifier trainer, these two adjectives are directly used to modify the image regions in this paper.

The region selection algorithm is shown in Table 1. Suppose m of N rectangular regions are to be selected. Here N is a huge number because even a small image can contain a very large quantity of possible regions[1]. Most of the N regions are by themselves with poor recognition performance. At first, according to the "accurate" criterion, the most accurate regions are selected. This can be done by training a classifier based on each region, calculating the recognition rate of the classifiers, and then choosing $n (m \ll n \ll N)$ regions with the highest recognition rates. Here n could be a given constant, or determined by the number of regions that achieve higher recognition rates than a given threshold θ. After that, according to the "diverse" criterion, these n regions are further selected based on their ability to rectify the errors made by the already-selected regions. This can be done by maintain a region set S, which initially contains only the region with the highest recognition rate. Then in each loop, the region with the best rectification ability is added into S. When there are m regions in S, the selection algorithm is over.

Obviously the algorithm shown in Table 1 is not the only way to select regions. As mentioned before, other feature selection approaches, such as the AdaBoost based selection similar to [9], are also applicable. In the later experimental part, the algorithm shown in Table 1 and the AdaBoost based selection will

[1]For instance, a 34×31 image may have $295,120$ possible rectangular regions

Table 1: Image region selection algorithm

- Given K example face images (each person has two images, one is gallery face and the other is probe face), m regions are to be selected from the exhausted set of N regions in the face images $R_1, R_2, \ldots, R_N$.
- For $i = 1 \ldots N$:
  1. Train a classifier based on the region R_i using the gallery faces as training set.
  2. Recognize all the probe faces using the classifier, getting the recognition rate of R_i, denoted by r_i .
- Sort R_i according to descending order of r_i. Get a sequence of regions: $R_{i1}, R_{i2}, \ldots, R_{iN}$.
- $S = \{R_{i1}\}$, $A = \{R_{i2}, \ldots, R_{in}\}$, where $m \ll n \ll N$.
- For $t = 1 \ldots (m-1)$:
  1. For each region R_i in A, calculate c_i, the number of the probe faces that R_i correctly recognizes but at least one of the regions in S doesn't.
  2. Find the region with the largest c_i, denoted by R_l.
  3. Remove R_l from A and add it to S.
- There are m selected regions in S.

be compared and analyzed in the framework of SEIR.

3 Selective Ensemble of Image Regions

After suitable Image regions for face recognition are selected, each of these regions is used to train a classifier. Then the classifiers are combined to recognize human faces. This is a general framework for a new category of face recognition algorithms, which can be called Selective Ensemble of Image Regions (abbreviated as SEIR). The flow chart of SEIR is shown in Figure 2.

There are two main steps in the training phase of SEIR. The first step is region selection. This can be done by the region selection algorithm shown in Table 1. Note that the classifier trainer in Table 1 is not specified. Any suitable classifier trainer for a certain problem can be adopted in the selection algorithm. Besides, more sophisticated region selection scheme other than the one shown in Table 1 may be applied in this step in the future. After region selection, the second step is classifier ensemble. Different classifier trainers and ensemble schemes may be adopted in this step. At last, the classifier ensemble is used in face recognition.

In this section, a realization of SEIR based on eigenspace, namely Selective Ensemble of Multiple Eigenspaces (abbreviated as SEME), is proposed.

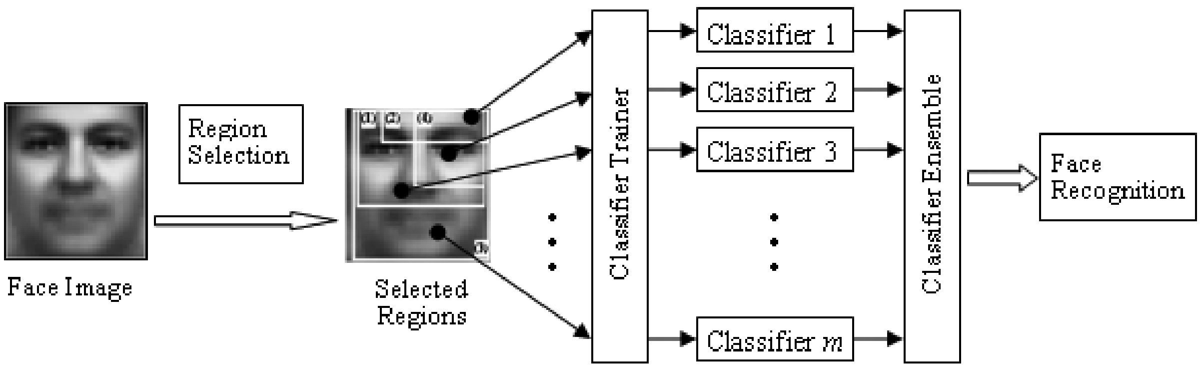
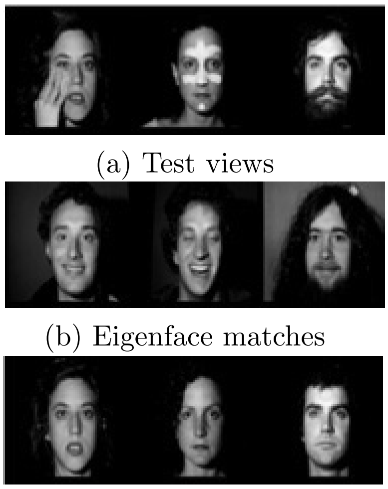


Figure 2: The flow chart of SEIR



(a) Test views

(b) Eigenface matches

(c) Eigenfeature matches

Figure 3: Examples that eigenface fails but eigenfeature succeeds [5]

3.1 Eigenspace Methods

Eigenspace methods have occupied a significant footing in face recognition techniques ever since the eigenface method [1] [2] was proposed. After more than ten years' development, eigenspace methods have formed an active, maybe the most active, algorithm family. The most representative members in this family may be eigenface [1] [2], eigenfeature [5], Bayesian matching [4][11][12] and PCA+LDA [13], among which eigenface and eigenfeature are closely related to SEME.

Eigenface is the application of a well-known technique in multivariate linear data analysis, PCA [3], in face recognition. The central idea of PCA is to reduce the dimensionality of a data set while retaining as much variation as possible. Eigenface regards each face image as a high dimensional vector by concatenating the intensity of each pixel in the image. The dimensionality n of the image vector $\mathbf{x}$ is usually very high. So PCA is used to reduce the dimensionality. It first constructs an eigenspace spanned by the $m(m \ll n)$ eigenvectors with the largest eigenvalues of the covariance matrix of the data set, and then project $\mathbf{x}$ into the m-dimensional eigenspace. The eigenvectors are called eigenfaces because their corresponding images look like human faces. The low-dimensional projection is then regarded as the feature vector of the face image and submitted to some classification methods, such as K-NN [14] [15], for face recognition.

Eigenfeature uses local facial features to construct multiple eigenspaces. Similar to eigenface, each local feature is projected into its corresponding eigenspace. The similarities between the corresponding features of two images are calculated according to their projections, and the similarities of all corresponding features are used to form a global similarity score of the two images. Face recognition is then performed based on the similarity score.

It is worth mentioning that Eigenfeature can be integrated with eigenface as an additional layer of description in terms of facial features [5]. This can be viewed as either a modular or layered representation of a face, where a coarse (low-resolution) description of the whole face is augmented by additional (high-resolution) details in terms of salient facial features [5]. A pure eigenface system might be fooled by gross variations in the input image (hats, beards, etc.), while the addition of the eigenfeature layer could help overcome this flaw. Figure 3 shows some examples where eigenface fails but eigenfeature succeeds. Moreover, in low-dimensional eigenspaces, eigenfeature outperforms eigenface in recognition [5]. These facts indicate that the combination of eigenface and eigenfeature can improve the recognition performance.

3.2 Selective Ensemble of Multiple Eigenspaces

SEME is a realization of SEIR, which uses the classifier trainer similar to that used in eigenface and eigenfeature. In fact, there are two classifier-training procedures in SEME. The first is in the region selection step, the other is in the classifier ensemble step. In both steps, SEME uses the same training scheme: first construct an eigenspace based on the image region, then project the region into the eigenspace to get the low dimensional feature vector and perform classification based on the feature vector. The similarity between two images on a particular region is defined as the reciprocal of the Euclidean distance between the corresponding feature vectors, and the similarity between two images is defined as the sum of similarities on all the selected regions. Given an unknown face image, its similarities to all the face images stored in the database are computed. Then the most similar one in the database is regarded as matching with the unknown face image.

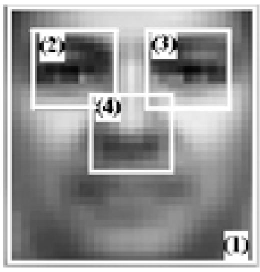
In the region selection step of SEME, one eigenspace must be constructed for each possible region in the image. This is a time-consuming procedure. It will be daunting to perform such a selection procedure each time when SEME is applied. Therefore the generalization of the selected regions, i.e. whether the regions selected on one data set can be directly used on other data sets, is crucial for using SEME in real applications. The generalization of the selected regions will be tested in the following experiments. If this property does exist, then the region selection procedure only need to be performed once, and then the selected regions will be stored for directly using in other applications.

It is worth mentioning that both eigenface and eigenfeature can be viewed as special cases of SEME. If in the region selection step, only one region, the whole face image is selected, then SEME is just eigenface. If in the region selection step, the regions around those salient facial features are selected, then SEME is just eigenfeature.

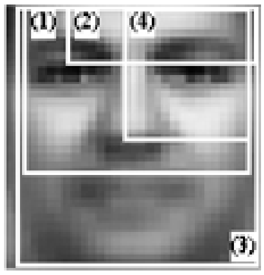
4 Experiments

4.1 Methodology

In the experiments, firstly the selection algorithm in Table1 is compared with the AdaBoost based selection in the framework of SEIR. Then the performance of SEME is compared with that of eigenface, PCA + LDA, eigenfeature, and the combination of eigenface and eigenfeature (denoted by eigenface + eigenfeature). After that the properties, the speedup, and the generalization ability of SEME are detailedly



(a) The rectangular features used by eigenface and eigenfeature



(b) The rectangular features selected by SEME

Figure 4: The rectangular features used for recognition

analyzed.

There are totally five face databases used in the experiments. The first is denoted by face database A, which is used to select rectangular regions and train eigenspaces. There are 400 gray-level frontal view images from 200 persons in the database. Each person has two images, one is used as gallery face and the other is used as probe face. The second database is denoted by face database B, which is used to test the performance of compared methods. There are 1,386 gray-level frontal view images from 693 persons in the database. Similarly, each person has two images, one is used as gallery face, and the other is used as probe face. All the images in face database A and B are randomly selected from the FERET face database [16]. The faces have significant variations in race, gender, age, expression, illumination, scale, etc. Note that there is no intersection between face databases A and B, i.e. the data set used to select rectangular regions and train eigenspaces is completely different from that used to test the algorithms. Thus the generalization of the selected regions can be tested. The third to fifth databases are all used to further test the generalization of the regions selected on face database A when applied on data sources other than FERET face database. The third database is denoted by face database C, the images in which are selected from AR face database [17]. Note that in the original AR face database, there are some faces with serious occlusions (sun glasses and scarf) and apparently different lighting sources. Since these variations are not included in face database A, which is used to select rectangular regions, all these faces are removed. There are 800 frontal view images from 100 persons in the database. Each person has eight images. Two are used as gallery faces, the other six are used as probe faces. The fourth database is denoted by face database D, the images in which are randomly selected from ORL face database [18]. There are 400 gray-level frontal view images from 40 persons in this database. Each person has ten images different in expression and illumination. Two are used as gallery faces and the other eight are used as probe faces. The fifth database is denoted by face database E, the images in which are randomly selected from BioID face database [19]. There are 132 gray-level frontal view images from 22 persons in the database. Each person has six images. two are used as gallery faces and the other four are used as probe faces.

All face images must be normalized. At first, the faces are cropped from the background to meet some constraints including that the line between the eyes is horizontal, the size of the image is fixed to 34×31 , and the center of the two eyes are fixed to the points $(9, 9)$ and $(9, 23)$. Then the cropped face images are histogram equalized. Note that accurately locating the eyes is critical to the normalization. Here the eyes are manually located.

The parameter n of the selection algorithm in Table 1 is set to 1000. As for the AdaBoost based selection, each subject in the training set is assigned a equal initial weight. After each it-

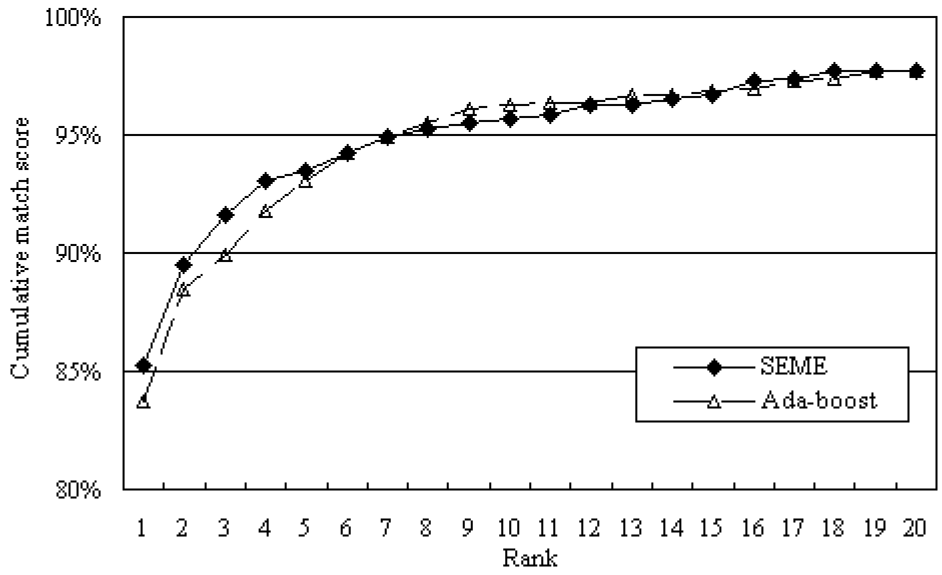


Figure 5: Comparison of SEME and AdaBoost based method on face database B

eration, one region is selected, and the weights are updated according to the recognition performance on that region. The similarity between two images is defined as the weighted sum of similarities on all the selected regions. The weight of each selected region is determined by its recognition performance on the training data. The rectangular regions used by eigenface and eigenfeature are shown in Figure 4 (a), where the rectangle (1) (the whole face) is used by eigenface, and the rectangles (2), (3) and (4) are used by eigenfeature. Since the difference between facial expressions is primarily articulated in the mouth, this feature is discarded for recognition purpose [5]. For convenience of observation, in Figure 4 the rectangular regions are superimposed on the mean face of face database B.

As mentioned in section 3.2, eigenface, eigenfeature and eigenface + eigenfeature can all be viewed as special cases of SEME. Therefore their training and testing process are intrinsically the same. PCA + LDA also goes in a similar way, except that an additional LDA is performed to get the final feature space. In all methods, during the training phase, the feature spaces are constructed on corresponding rectangular regions. During the test phase, each rectangular region is projected into the corresponding feature space to get its feature vector. In the experiments, the dimension of the eigenspace is set to 20 (including the first step of PCA + LDA) if not explicitly stated. The similarity between two images is defined as that used in SEME, which is described in section 3.2. For each probe face, the similarities between it and every gallery face are computed. Then the gallery faces are ranked according to the descending order of the similarities. The identity of the top k images in the list is considered as the recognition result. Note that conforming to the FERET testing protocol [16], both "*is the top match correct?*" and "*is the correct answer in the top k matches?*" are considered.

4.2 Results

At first, four rectangular regions are selected by the algorithm shown in Table1 on face database A. The selected regions are shown in Figure 4 (b), where the number of a particular rectangle corresponds to the order that it is selected. It can be seen that the selected regions are similar to those used by eigenface and eigenfeature in the following aspects:

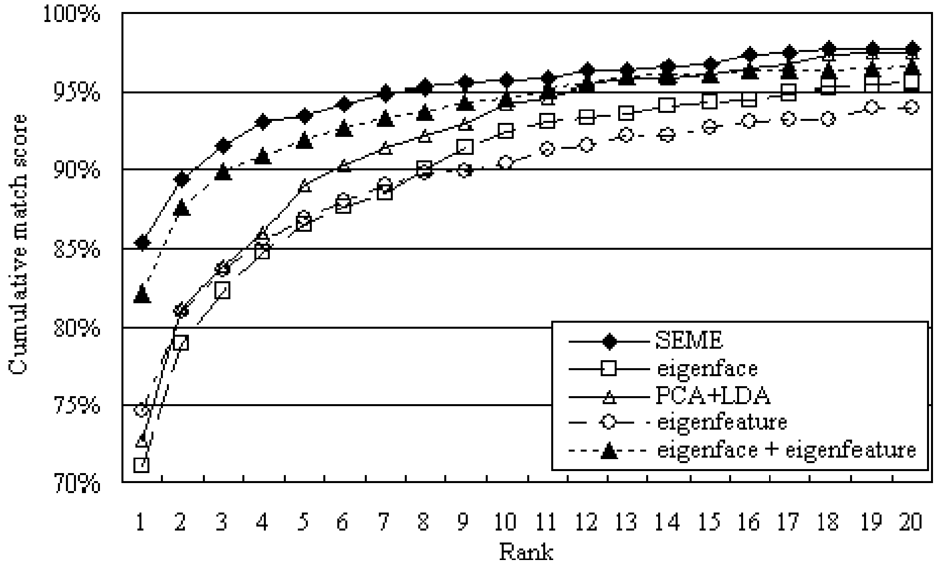


Figure 6: Comparison of SEME, eigenface, PCA + LDA, eigenfeature, and eigenface + eigenfeature on face database B

1. The locations of most rectangles are around eyes and nose. This is consistent with the intuition that these facial features are critical in face recognition.

2. The selected rectangle (3) is almost the whole face. This illuminates the indispensability of the global feature in face recognition.

3. Except for the global feature (3), all the selected features do not include the mouth. This supports the claim that the difference between facial expressions is primarily articulated in the mouth [5].

However, the selected regions have the following differences from those used by eigenface and eigenfeature:

1. Eyes and nose are all included in rectangle (1), rather than distributed in several rectangles like eigenfeature.

2. Rectangle (2) introduces a new facial feature, i.e. eyebrows. While in Figure 4 (a), the eyebrows are included in the rectangles of eyes.

3. Some rectangles are not symmetric about the midline of the face image. The reason may be that human faces are usually symmetric so that a slightly more than half of the face may contain almost all information.

4. Rectangle (4) contains an eye and a part of the nose. This indicates that the rectangular regions are not necessary to use a single facial feature as basic unit.

AdaBoost is also used to select four regions. After that, the recognition performances of them and those shown in Figure 4 (b) are compared on face database B. The result is shown in Figure 5, where the horizontal axis is rank and the vertical axis is cumulative match score. When rank $= k (k = 1 \ldots 20)$, the value of cumulative match score means the probability of that the correct result is in the top k matches. It can be seen that SEME performs better than the AdaBoost based method in most cases. This might because that the AdaBoost method tends to overfit the training set, especially when training on a relatively small data set (face database A) and test on a large data set (face database B). Moreover, the

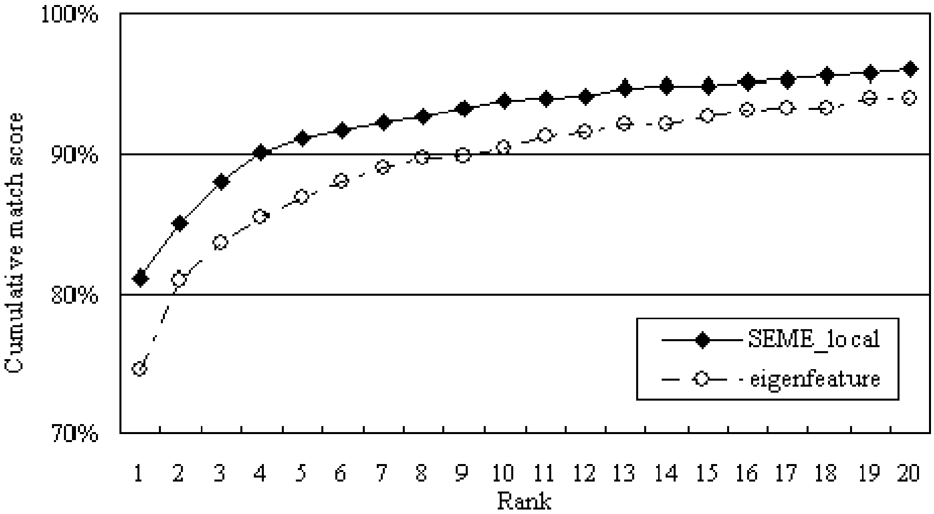
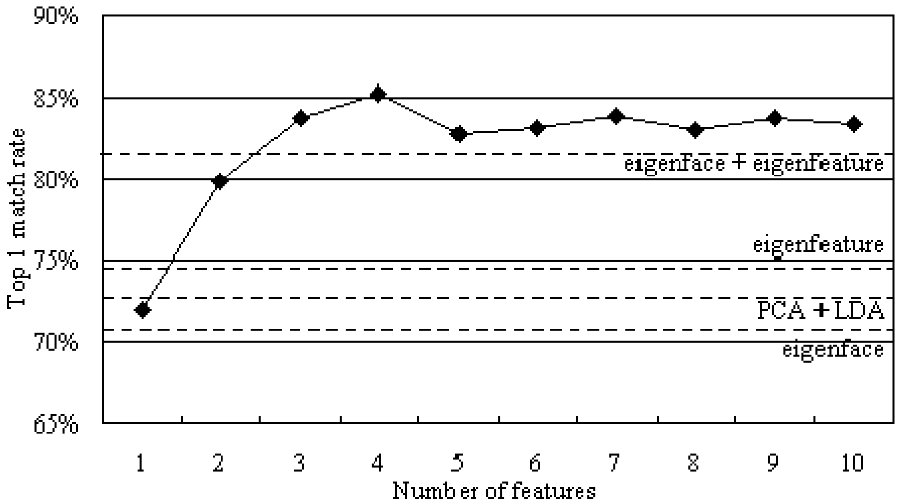


Figure 7: Comparison of SEME_ local and eigenfeature on face database B

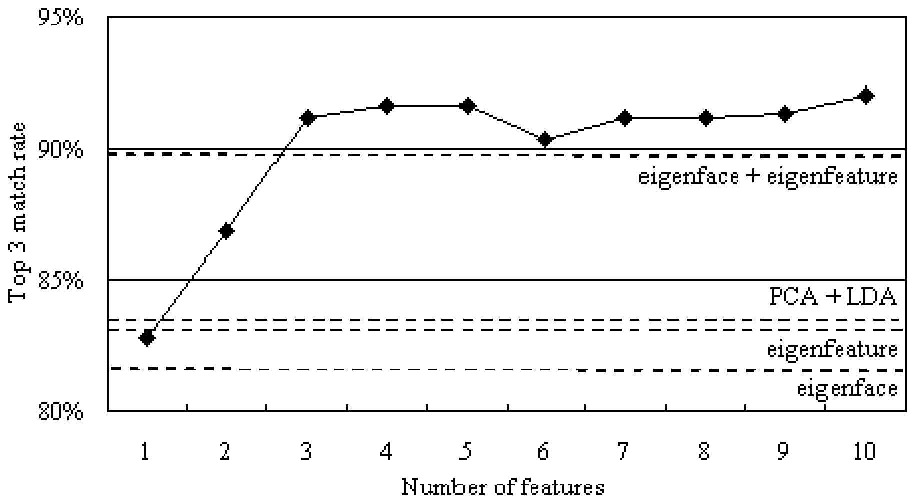
training of SEME only needs to scan all possible regions (i.e. to train a classifier on each region and test it) once no matter how many regions are to be selected. However, the AdaBoost selection must scan all regions once before each region can be selected. Thus in the configuration of this experiment (four regions), the training procedure of AdaBoost selection is about four times as long as that of SEME. As mentioned before, the amount of possible rectangular regions in a image is very huge so that the training of the AdaBoost based method is very time-consuming. Since SEME is better and faster than the AdaBoost based method, in the following experiments, only SEME is compared with other existing techniques and further analyzed.

The recognition performances of SEME, eigenface, PCA + LDA, eigenfeature and eigenface + eigenfeature on face database B are shown in Figure 6, which shows that SEME achieves the best performance. Since both SEME and eigenface + eigenfeature use four rectangles, comparison between them is more meaningful. As can be seen, the cumulative match score of SEME is about 2% to 3% higher than that of eigenface + eigenfeature. It can also be found that although eigenfeature is better than eigenface when rank is smaller than 8, it is gradually surpassed by eigenface with the increase of rank. As reported in previous works, PCA + LDA always performs better than eigenface. It is also better than eigenfeature in most cases, but worse than eigenface + eigenfeature. Since eigenface + eigenfeature combines both the global feature and local features, its performance is better than PCA + LDA, eigenfeature and eigenface, but still worse than SEME.

As can be seen in Figure 4 (b), the rectangle (3) is almost the whole face image. So it can be viewed as a global feature and the other three rectangles are local features. To compare the performance of these local features and those used by eigenfeature, SEME using only rectangle (1), (2) and (4) (denoted by SEME_ local) is tested on face database B. The results, together with those of eigenfeature, are shown in Figure 7. Figure 7 reveals that SEME_ local outperforms eigenfeature even without the global feature. Since the rectangle (3) used by SEME is a global feature similar to the one used by eigenface, we believe that the superiority of SEME over eigenface + eigenfeature mainly comes from the superiority of SEME_ local over eigenfeature, i.e. the superiority of

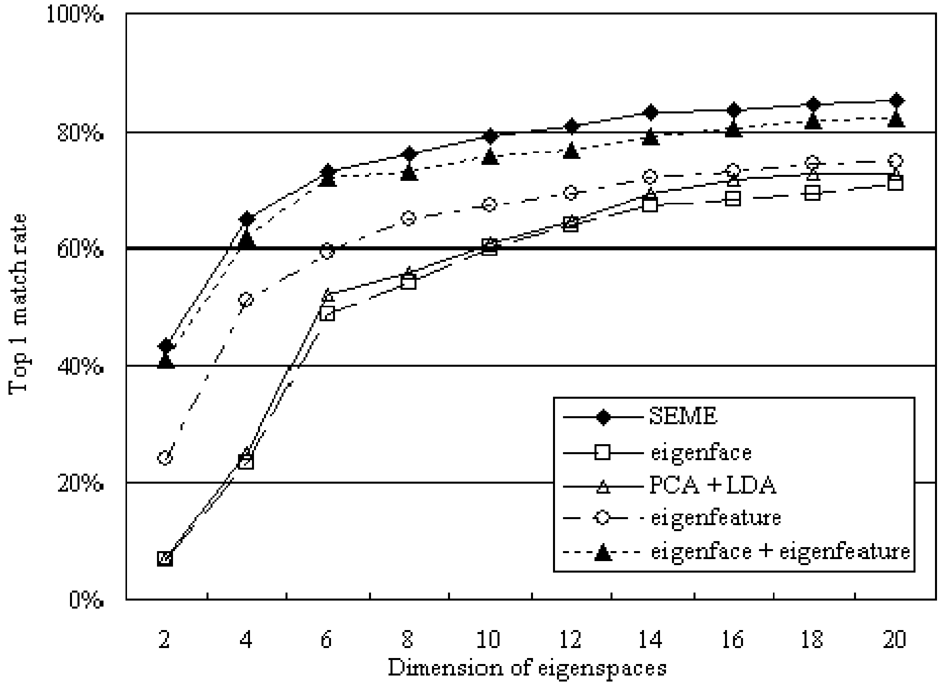


(a) Top 1 match rate of SEME with different number of selected features

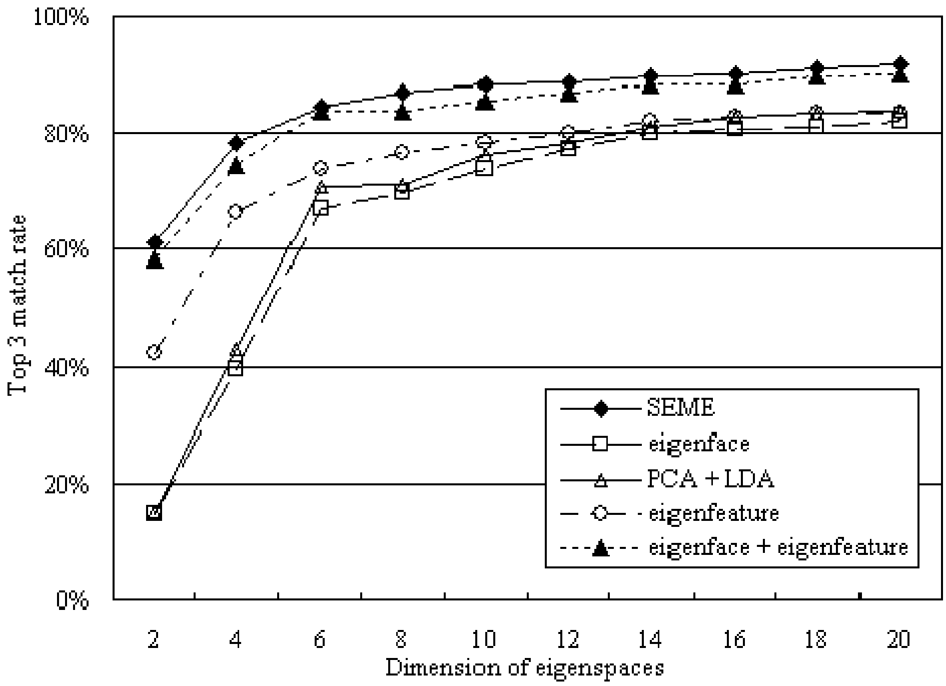


(b) Top 3 match rate of SEME with different number of selected features

Figure 8: Performance of SEME with different number of selected features



(a) Top 1 match rate with different dimensions of eigenspaces



(b) Top 3 match rate with different dimensions of eigenspaces

Figure 9: Comparison under different dimensions of eigenspaces

the selected local features over empirically determined ones.

4.3 Properties of SEME

Figure 4 (b) only shows the first four selected rectangles. In fact, there is no limitation to the amount of the rectangular regions used in SEME. To investigate the properties of SEME when the number of selected rectangular regions increases, SEME with different number of regions (from 1 to 10 regions) are tested. The results are shown in Figure 8, where Figure 8(a) shows the top 1 match rate and Figure 8(b) shows the top 3 match rate. For the convenience of comparison, the performances of eigenface, PCA + LDA, eigenfeature and eigenface + eigenfeature are shown as four dashed lines in the figures.

Figure 8 reveals that with the increase of the amount of regions, the match rate of SEME gradually goes up and surpasses those of other compared methods. SEME achieves the best performance when four regions are used. After that, the match rate of SEME begins to fluctuate slightly (in a range of about 1%), but still keeps better than all the other methods.

In the above experiments, the dimensions of eigenspaces are all fixed to 20. In order to further analyze the properties of SEME with different dimensions of eigenspaces, the top 1 match rate and top 3 match rate of SEME, eigenface, PCA + LDA, eigenfeature, and eigenface + eigenfeature are compared as the dimension of eigenspaces increases gradually from 2 to 20 with 2 as the interval. The results are shown in Figure 9.

Figure 9 reveals that SEME achieves the best performance among the compared methods, no matter top 1 match rate or top 3 match rate is concerned. In addition, Figure 9 supports the claim of Pentland et al. [5], i.e. in the lower dimensions of eigenspaces, eigenfeature remarkably outperforms eigenface. But with the increase of the dimensionality, this superiority gradually shrinks. Moreover, when the dimensionality is low (2 to 4), the performance of PCA + LDA is very close to that of eigenface. This is because the low dimensional eigenspace loses too much information in the original data set so that LDA can not efficiently find a discriminative space. However, with higher dimension, PCA + LDA keeps a steady superiority over eigenface.

4.4 Exploit Symmetry to Speedup Region Selection

The region selection algorithm needs to search all possible regions in the face image, so it is computational expensive. Taking into consideration that human faces are approximately symmetric, the region selection procedure can be significantly accelerated through searching only half of the face image and then obtaining the rectangles in the other half by symmetry. Here one feature comprises two rectangles that are symmetric about the midline of the face image. The vector of this kind of feature is obtained by concatenating the vectors of these two rectangles. This variation of SEME is denoted by SEME_ sym. The first four features selected by SEME_ sym on face database A are shown in Figure 10, where the numbers under the images correspond to the selection order. It is easy to find out that these features are similar with those in Figure 4 (b). For example, they are all selected in the following order: (1) eyes and nose, (2) eyebrows, (3) the whole face, (4) eyes.

The performance of SEME_ sym is tested on face database B. The results are shown in Figure 11, which reveals that although SEME_ sym is worse than SEME, it is still better than eigenface + eigenfeature. Thus when training time is under serious consideration, SEME_ sym will be a good option.

4.5 The Generalization of the Selected Regions

Since the computational cost of region selection is expensive, the generalization ability of the selected regions is crucial for SEME to be

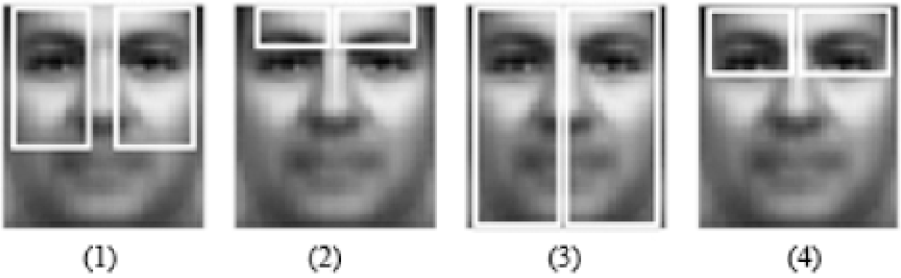


Figure 10: The symmetric rectangular features selected by SEME_ sym

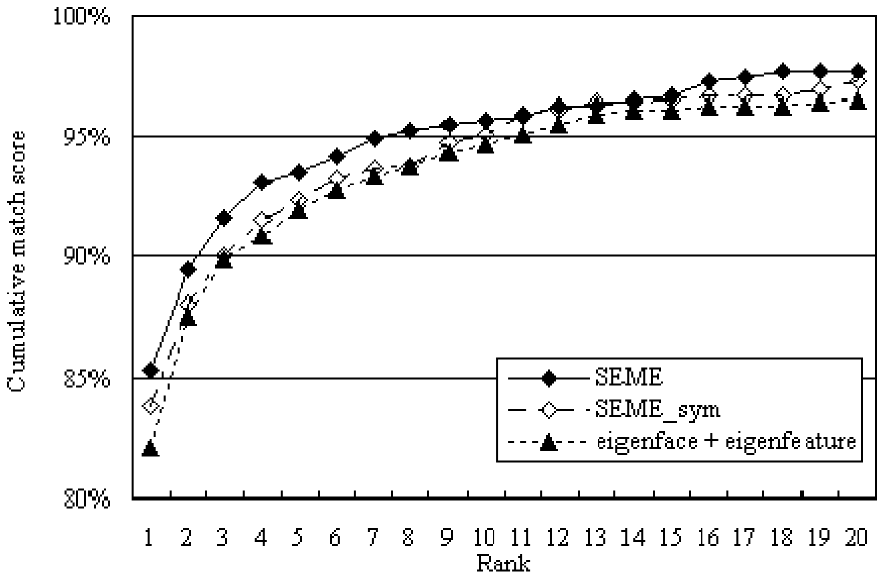


Figure 11: Performance of SEME_ sym compared with SEME and eigenface + eigenfeature

used in real applications. In the former experiments, the data set used to select the regions and the data set used to test the algorithms are completely different. Therefore the good performance of SEME illuminates the good generalization ability of the selected regions. But after all, face database A and face database B are all from the FERET database. In this section, the generalization ability of the regions selected on face database A will be further tested on face database C (subset of AR face database [17]), D (ORL face database [18]) and E (subset of BioID face database [19]).

The recognition performances of SEME, eigenface, PCA + LDA, eigenfeature and eigenface + eigenfeature on face database C are shown in Figure 12. Figure 12 reveals that SEME using the regions selected on face database A achieves the best performance on face database C. It can be seen that Figure 12 is very similar to Figure 6. Except for SEME, eigenface + eigenfeature achieves the best performance and then is PCA + LDA. Eigenfeature is comparable with eigenface in the lower ranks, but with the increase of rank, it is greatly surpassed by eigenface. From these two figures, we can conclude that the five compared face recognition techniques, at least under the configurations of this experiment, can be sorted according to their performance as: SEME > eigenface + eigenfeature > PCA + LDA > eigenface > eigenfeature.

The recognition performances of SEME, eigenface, PCA + LDA, eigenfeature and eigen-

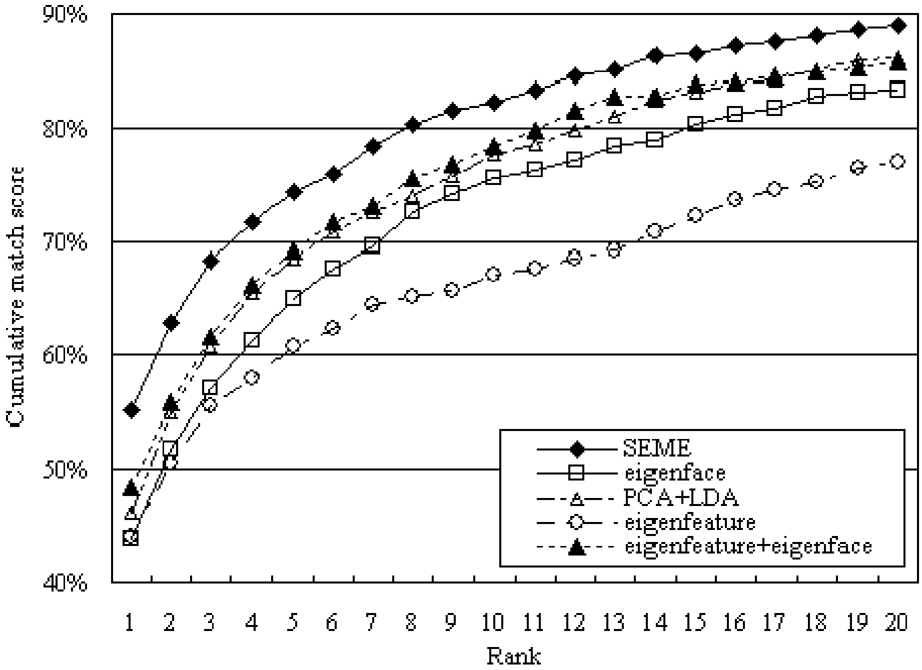


Figure 12: Comparison of SEME, eigenface, PCA + LDA, eigenfeature, and eigenface + eigenfeature on face database C

face + eigenfeature on face database D are shown in Figure 13. Note that since there are only 40 persons in face database D, the maximum value of rank is reduced to 5. Figure 13 reveals that SEME using the features selected on face database A achieves the best performance on face database D. Both Figure 12 and Figure 13 indicate that even on different data sources, SEME still has good generalization ability.

There are two main differences between Figure 13 and Figure 6. The first is that PCA + LDA performs better than eigenface + eigenfeature. This might be because that with fewer classes (only 40 different persons) in the database, the procedure of LDA can find a better discriminative subspace. The second difference is that the performance of eigenface is better than that of eigenfeature with any value of rank. A possible explanation to this is that since there are fewer classes in the database, the variation of different patterns is limited. Thus the global feature alone is sufficient in achieving a high recognition rate. In order to verify these explanations, experiments on face database E, a database of even fewer classes (only 22 persons), are performed. The results are shown in Figure 14. It can be seen that since there are too few persons in the database, all of SEME, eigenface and PCA + LDA can correctly recognize all the probe faces even when rank = 1. On the other hand, the recognition rate of eigenfeature and eigenface + eigenfeature cannot reach 100% even when the value of rank increases to 5. This illuminates that when the number of different persons (classes) in the database is relatively small, the global feature performs better than the local features. But even in this situation, SEME is still a good choice.

5 Conclusion

In this paper, all the possible rectangular regions in the face image are regarded as a certain kind of features. From this point of view, face recognition techniques can be divided into two steps: The fist step is to select one or

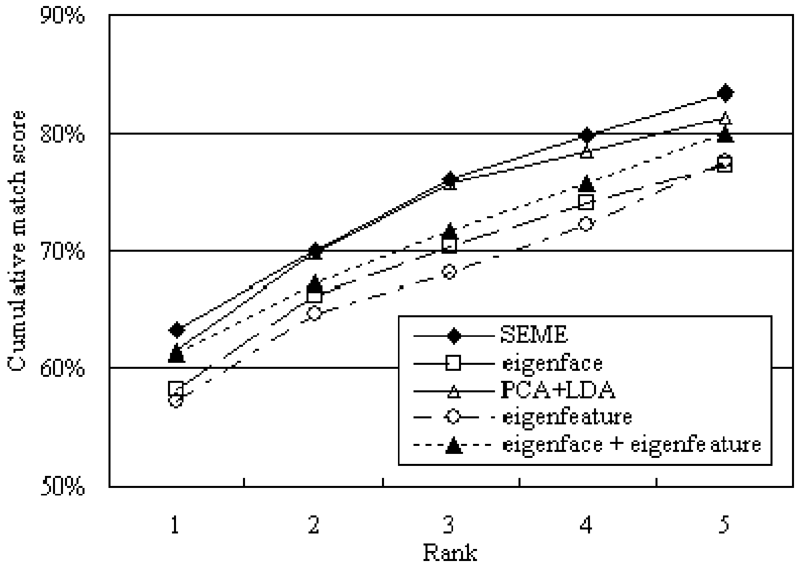


Figure 13: Comparison of SEME, eigenface, PCA + LDA, eigenfeature, and eigenface + eigenfeature on face database D

more features from all possible feature candidates, the second step is to train a classifier for face recognition. Currently almost all face recognition techniques only focus on the second step. The region selection in the first step is empirically performed by the operators, consciously or unconsciously. In this paper, an automatic region selection algorithm for face recognition is proposed. Based on this selection algorithm, a novel framework for face recognition, namely SEIR, is proposed. A realization of SEIR named SEME is then designed based on multiple eigenspaces. SEME is tested and compared with eigenface, PCA + LDA, eigenfeature and eigenface + eigenfeature. The experimental results show that SEME achieves the best performance.

The computational cost of the selection algorithm is expensive because all possible rectangular regions in the face images must be considered. However, the symmetry of human faces can be utilized to speedup the procedure, which could significantly reduce the computational cost. Moreover, the selected regions have good generalization ability, i.e. once the regions are selected, they can be directly used on other data sets. It can be imagined that we can use a powerful computer to run the region selection algorithm on very large quantity of face images. Once the region set is selected, it will be distributed for various applications of face recognition.

The main contribution of this paper is a new framework named SEIR for face recognition. SEME is only a intuitive realization in this framework, and is used to illuminate the usefulness of SEIR. Although SEME has been compared with the AdaBoost based method and proved to be better, other more efficient selection methods might be developed for SEIR in the future. There still leaves a lot of work to do along this approach.

SEME uses eigenspace to train classifiers. Other classifier trainers for face recognition, such as Bayes matching [4][11][12] and PCA+LDA [13], can also be integrated in the framework of SEIR. The performance of these variations needs to be further tested.

The rectangular region can be extended to some more complex regions, such as regions that comprise several rectangles. It will be interesting to explore in the future that whether more complex features can bring benefits to face recognition.

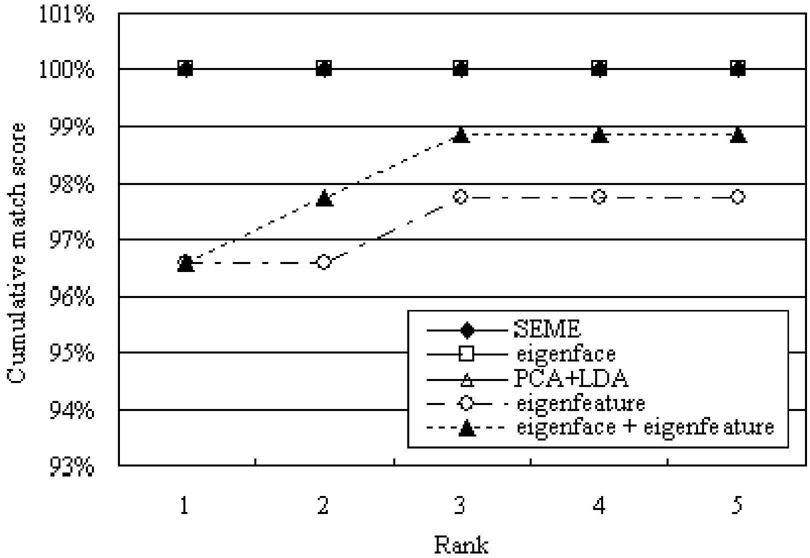


Figure 14: Comparison of SEME, eigenface, PCA + LDA, eigenfeature, and eigenface + eigenfeature on face database E

According to Brunelli and Poggio [6], the combination of the similarity scores of different features can be done in the following ways:

- Choose the score of the most similar feature.
- The feature scores are added.
- The feature scores are added, but each feature is given a different weight (the same for all people).
- The features are added using a person-dependent weight.
- The feature scores are used as inputs to a classifier such as a nearest neighbor or a HyperBF network [20].

The strategy adopted in SEME is the second one, i.e. all the feature scores are simply added. Other strategies are worth to be studied in the future.

Acknowledgements

The authors want to thank Dr. Aleix M. Martinez for providing the AR face database. Portions of the research in this paper use the FERET database of facial images collected under the FERET program.

References

[1] Turk M, Pentland A. Eigenfaces for recognition. Journal of Cognitive Neuroscience, 1991, 3(1): 71-86.

[2] Moghaddam B, Pentland A. Probabilistic visual learning for object representation. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1997, 19(7): 696-710.

[3] Jolliffe I T. Principal Component Analysis. New York: Springer-Verlag, 1986.

[4] Moghaddam B, Jebara T, Pentland A. Bayesian face recognition. Pattern Recognition, 2000, 33(11): 1771-1782.

[5] Pentland A, Moghaddam B, Starner T. View-based and modular eigenspaces for face recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Seattle, WA, 1994, pp. 84-91.

[6] Brunelli R, Poggio T. Face recognition: feature versus templates. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1993, 15(10): 1042-1052.

[7] Zhou Z-H, Wu J, Tang W. Ensembling neural networks: many could be better than all. Artificial Intelligence, 2002, 137(1-2): 239-263.

[8] Penev P S, Atick J J. Local feature analysis: a general statistical theory for object representation. Network: Computation in Neural Systems, 1996, 7: 477-500.

[9] Yang P, Shan S, Gao W, Li S Z, Zhang D. Face recognition using Ada-Boosted Gabor features. In: Proceedings of the 6th International Conference on Automatic Face and Gesture Recognition. Seoul, Korea, 2004, pp. 356-361.

[10] Krogh A, Vedelsby J. Neural network ensembles, cross validation, and active learning. In: Tesauro G, Touretzky D, Leen T, eds., Advances in Neural Information Processing Systems 7, Cambridge, MA: MIT press, 1995, pp. 231-238.

[11] Moghaddam B, Wahid W, Pentland A. Beyond eigenfaces: probabilistic matching for face recognition. In: Proceedings of the IEEE International Conference on Automatic Face and Gesture Recognition, Nara, Japan, 1998, pp. 30-35.

[12] Moghaddam B. Principal manifolds and probabilistic subspaces for visual recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002, 24(6): 780-788.

[13] Zhao W, Chellappa R, Philips P. Subspace Linear Discriminant Analysis for Face Recognition. Technical Report CAR-TR-914. Center for Automation Research, University of Maryland, 1999.

[14] Ho T K. Nearest Neighbors in random subspaces. In: Amin A, Dori D, Pudil P, Freeman H, eds., Lecture Notes in Computer Science 1451, Berlin: Springer, 1998, pp.640-648.

[15] Lowe D. Similarity metric learning for a variable-kernel classifier. Neural Computation, 1995, 7(1): 72-85.

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

[17] Martinez A M, Benavente R. The AR Face Database. CVC Technical Report #24, June 1998

[18] Samaria F, Harter A. Parameterisation of a stochastic model for human face identification. In: Proceedings of the 2nd IEEE International Workshop on Applications of Computer Vision, Sarasota, FL, 1994, pp. 138-142.

[19] The BioID Face Database, http://www.bioid.com/research/index.htm, BioID-Technology Research, June 2001.

[20] Poggio T, Girosi F. A theory of networks for recognize three-dimensional objects. Nature, 1990, 343 (6225): 1-3.