COMBINING CLASSIFIERS FOR FACE RECOGNITION

Xiaoguang Lu*, *Yunhong Wang*†, *Anil K. Jain**

*Department of Computer Science & Engineering, Michigan State University
East Lansing, MI 48824
{lvxiaogu, jain}@cse.msu.edu

†National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences
P.O. Box 2728, Beijing 100080, P. R. China
wangyh@nlpr.ia.ac.cn

ABSTRACT

Current two-dimensional face recognition approaches can obtain a good performance only under constrained environments. However, in the real applications, face appearance changes significantly due to different illumination, pose, and expression. Face recognizers based on different representations of the input face images have different sensitivity to these variations. Therefore, a combination of different face classifiers which can integrate the complementary information should lead to improved classification accuracy. We use the sum rule and RBF-based integration strategies to combine three commonly used face classifiers based on PCA, ICA and LDA representations. Experiments conducted on a face database containing 206 subjects (2,060 face images) show that the proposed classifier combination approaches outperform individual classifiers.

1. INTRODUCTION

Human face recognition has a tremendous potential in a wide variety of commercial and law enforcement applications. Considerable research efforts have been devoted to the face recognition problem over the past decade [1]. Although there are a number of face recognition algorithms which work well in constrained environments, face recognition is still an open and very challenging problem in real applications.

Among face recognition algorithms, appearance-based approaches [2][3][4][5] are the most popular. These approaches utilize the pixel intensity or intensity-derived features. Several such systems have been successfully developed and installed [1][6][7][8]. However, appearance-based methods do not perform well in many real-world situations, where the query test face appearance is significantly different from the training face data, due to variations in pose, lighting and expression. Some examples of these variations for one of the subjects in our database are illustrated in Fig. 1. While a robust classifier could be designed to handle any one of these variations, it is extremely difficult for an appearance-based approach to deal with all of these variations. Each individual classifier has different sensitivity to different changes in the facial appearance. It has been reported that each appearance-based method shows different levels of performance on different subsets of images [6], suggesting that different classifiers contribute complementary information to the classification task. A combination scheme involving different face classifiers, which integrates various information sources, is likely to improve the overall system performance.

Fig. 1. Facial variations under different lighting conditions and facial expressions for the same subject [9].

The classifier combination can be implemented at two levels, feature level and decision level. We use the decision level combination that is more appropriate when the component classifiers use different types of features. Kittler [10] provides a theoretical framework to combine various classifiers at the decision level. Many practical applications of combining multiple classifiers have been developed. Brunelli and Falavigna [11] presented a person identification system by combining outputs from classifiers based on audio and visual cues. Jain et al. [12] integrated multiple fingerprint matchers to develop a robust fingerprint verification system. Hong and Jain [13] designed a decision fusion scheme to combine faces and fingerprint for personal identification. Marcialis and Roli [14] exploited the fusion of

This research was supported by NSF IUC on Biometrics (CITeR), at West Virginia University.

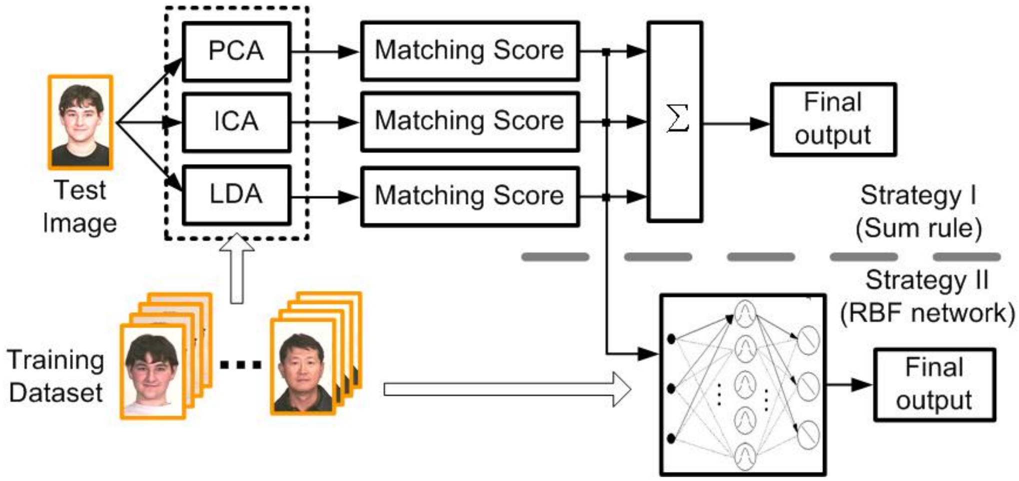

Fig. 2. Classifier combination system framework.

PCA and LDA for face verification.

We propose two combination strategies, sum rule and RBF network, to integrate the outputs of three well-known appearance-based face recognition methods, namely PCA [2], ICA [3] and LDA [4][5]. Our combination strategy is designed at the decision level, utilizing all the available information, i.e. a subset of (face) labels along with a confidence value, called the matching score provided by each of the three face recognition method.

2. CLASSIFIER INTEGRATION

Our combination scheme is illustrated in Fig. 2. While this framework does not limit the number of component classifiers, we currently use only three classifiers, namely, PCA, ICA and LDA. Following two strategies are provided for integrating outputs of individual classifiers, (i) the sum rule, and (ii) a RBF network as a classifier, using matching scores as the input feature vectors.

2.1. Appearance-based Face Classifiers

Three appearance-based classifiers, PCA [2], ICA [3] and LDA [4][5] have been implemented. In each of these approaches, the 2-dimensional face image is considered as a vector, by concatenating each row (or column) of the image. Each classifier has its own representation (basis vectors) of a high dimensional face vector space. By projecting the face vector to the basis vectors, the projection coefficients are used as the feature representation of each face image. The matching score between the test face image and training data is calculated as the cosine value of the angle between their coefficients vectors.

Let $X = (x_1, x_2, \ldots, x_i, \ldots, x_N)$ represent the $n \times N$ data matrix, where each x_i is a face vector of dimension n, concatenated from a $p{\times}p$ face image, where $p{\times}p = n$. Here n represents the total number of pixels in the face image and N is the number of face images in the training set. The mean vector of the training images $\mu = \sum_{i=1}^{N} X_i$ is subtracted from each image vector. All the three representations can be considered as a linear transformation from the original image vector to a projection feature vector, i.e.

$$Y = W^T X, \tag{1}$$

where Y is the $d \times N$ feature vector matrix, d is the dimension of the feature vector, and W is the transformation matrix. Note that $d << n$.

1. PCA [2]. The Principal Component Analysis basis vectors are defined as the eigenvectors of the scatter matrix S_T,

   $$S_T = \sum_{i=1}^{N} (x_i - \mu)(x_i - \mu)^T. \tag{2}$$

   The transformation matrix W_{PCA} is composed of the eigenvectors corresponding to the d largest eigenvalues. After applying the projection, the input vector (face) in an n-dimensional space is reduced to a feature vector in a d-dimensional subspace.

2. ICA [3]. Bartlett et al. [3] provided two architectures based on Independent Component Analysis, statistically independent basis images and a factorial code representation, for the face recognition task. The ICA separates the high-order moments of the input in addition to the second-order moments utilized in PCA. Both the architectures lead to a similar performance. There is no special order imposed on the ICA basis vectors.

3. LDA [4][5]. The Linear Discriminant Analysis finds a transform W_{LDA}, such that

$$W_{LDA} = \arg\max_{W} \frac{W^T S_B W}{W^T S_W W}, \tag{3}$$

where S_B is the between-class scatter matrix and S_W is the within-class scatter matrix, defined as

$$S_B = \sum_{i=1}^{c} N_i (x_i - \mu)(x_i - \mu)^T, \tag{4}$$

$$S_W = \sum_{i=1}^{c} \sum_{x_k \in X_i} (x_k - \mu_i)(x_k - \mu_i)^T. \tag{5}$$

In the above expression, N_i is the number of training samples in class i, c is the number of distinct classes, μ_i is the mean vector of samples belonging to class i and X_i represents the set of samples belonging to class i.

2.2. Integration Strategy

Kittler [10] analyzed several classifier combination rules and concluded that the sum rule (defined below) outperforms other combination schemes based on empirical observations. Unlike explicitly setting up combination rules, it is possible to design a new classifier using the outputs of individual classifiers as features to this new classifier. We adopt the RBF network [15] as this new classifier. Given m templates in the training set, m matching scores will be output for each test image from each classifier. We consider the following two integration strategies

1. Strategy I: Sum Rule. The combined matching score is calculated as

$$MS_{comb} = MS_{PCA} + MS_{ICA} + MS_{LDA}. \tag{6}$$

For a given test sample, Output the class with the largest value of MS_{comb}.

2. Strategy II: RBF network. For each test image, the m matching scores obtained from each classifier are used as a feature vector. Concatenating these feature vectors derived from three classifiers results in a feature vector of size $3m$. An RBF network is designed to use this new feature vector as the input to generate classification results. We adopt a 3-layer RBF network. The input layer has $3m$ nodes and the output has c nodes, where c is the total number of classes (number of distinct faces). In the output layer, the class corresponding to the node with the maximum output is assigned to the input image. The number of nodes in the hidden layer is constructed empirically, depending on the sizes of the input and output layers.

3. EXPERIMENTS AND DISCUSSION

Our database is a collection of four different face databases, available in the public domain (see table 1). There are 206 subjects with 10 images per subject for a total of 2,060 images. Face images selected are near frontal and contain variations in pose, illumination and expression. Some images in the individual databases are not selected for our experiments; these face images have out-of-plane rotation by more than 45 degrees in the NLPR+MSU database and face images with occlusions due to sun glasses or a scarf in the AR database. Sample images from the databases are shown in Fig. 3. Face images are closely cropped to include only the internal facial structures such as the eyebrows, eyes, nose and mouth, and aligned by the centers of the two eyes. All cropped images are resized to 42×42. Each image vector is normalized to be of unit length.

Table 1. Database description.

Face database	No. of subjects	Variations included
ORL [16]	40	Slight pose and expression
Yale [9]	15	Illumination and expression
AR [17]	120	Illumination and expression
NLPR+MSU (collected by the authors)	31	Slight pose and expression

Fig. 3. Representative face images in the database. (a) ORL, (b) Yale, (c) AR and (d) NLPR+MSU.

The entire face database is divided into two parts. Nine images of each subject are used to construct the training data and the remaining one is used for testing. This partition is repeated 10 different times so that every image of the subject can be used for testing. The classification accuracy is the average of these ten different tests.

All the individual classifiers use the cosine value of the angle between the two projection coefficient vectors (one from the database image and the other from the test image) as the matching score. Database image with the best match is used to determine the classification of the input image. The sum rule is applied to the matching score outputs of the three classifiers. The database image with the maximum sum score is output as the final result. The recognition accuracies of different face recognition approaches are listed in table 2. The cumulative match score vs. rank curve [6] is used to show the performance of each classifier, see Fig. 4. Since our RBF network outputs the final label, no rank information is available. As a result, we cannot compute the cumulative match score vs. rank curve for RBF combination.

Table 2. Recognition accuracy of different classifiers.

PCA	ICA	LDA	Sum rule	RBF based
79.1%	88.1%	81.0%	90.0%	90.2%

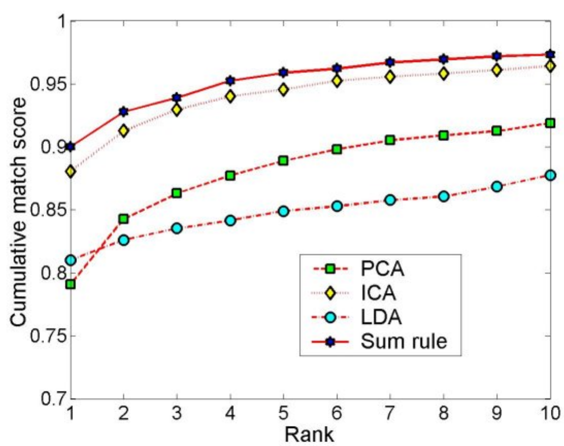

Fig. 4. *Cumulative match score vs. rank* curve for the sum rule.

Table 2 and figure 4 show that the combined classifiers, based on both the sum-rule and RBF network, outperform each individual classifier.

4. CONCLUSIONS AND FUTURE WORK

An integration scheme, which combines the output matching scores of three well-known face recognition approaches, is proposed to improve the performance of a face identification system. Two combination strategies, sum rule and RBF-based integration, are implemented to combine the output information of three individual classifiers, namely PCA, ICA and LDA. The proposed system framework is scalable; other face recognition modules can be easily added into this framework. Experimental results are encouraging, illustrating that both the combination strategies lead to more accurate face recognition than that made by any one of the individual classifiers. We are currently investigating the weighted sum rule based on the user-specific matching score distribution.

5. REFERENCES

[1] W. Zhao, R. Chellappa, A. Rosenfeld, and P.J. Phillips, “Face recognition: A literature survey,” *CVL Technical Report, University of Maryland*, Oct. 2000, <ftp://ftp.cfar.umd.edu/TRs/CVL-Reports-2000/TR4167-zhao.ps.gz>.

[2] M. Turk and A. Pentland, “Eigenfaces for recognition,” *Journal of Cognitive Neuroscience*, vol. 3, no. 1, pp. 71–86, Mar. 1991.

[3] M.S. Bartlett, H.M. Lades, and T.J. Sejnowski, “Independent component representations for face recognition,” in *Proceedings of the SPIE*, 1998, vol. 3299, pp. 528–539.

[4] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, “Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection,” *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, pp. 711–720, Jul. 1997.

[5] D. L. Swets and J. Weng, “Using discriminant eigenfeatures for image retrieval,” *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 18, no. 8, pp. 831–836, 1996.

[6] P. Jonathon Phillips, Hyeonjoon Moon, Syed A. Rizvi, and Patrick J. Rauss, “The feret evaluation methodology for face-recognition algorithms,” *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 22, no. 10, pp. 1090–1104, 2000.

[7] Identix, <*http://www.identix.com/*>, Minnetonka, MN.

[8] Viisage, <*http://www.viisage.com/*>, Littleton, MA.

[9] *Yale University face database*, <http://cvc.yale.edu/projects/yalefaces/yalefaces.html>.

[10] J. Kittler, M. Hatef, R. Duin, and J. Matas, “On combining classifiers,” *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 20, no. 3, pp. 226–239, 1998.

[11] R. Brunelli and D. Falavigna, “Person identification using multiple cues,” *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 17, no. 10, pp. 955–966, Oct. 1995.

[12] A. K. Jain, S. Prabhakar, and S. Chen, “Combining multiple matchers for a high security fingerprint verification system,” *Pattern Recognition Letters*, vol. 20, no. 11-13, pp. 1371–1379, 1999.

[13] L. Hong and A.K Jain, “Integrating faces and fingerprint for personal identification,” *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 20, no. 12, pp. 1295–1307, 1998.

[14] G. L. Marcialis and F. Roli, “Fusion of lda and pca for face verification,” in *Biometric Authentication*, Jun. 2002, vol. LNCS 2359, pp. 30–37.

[15] C. M. Bishop, *Neural Networks for Pattern Recognition*, Oxford University Press, UK, 1995.

[16] Ferdinando Samaria and Andy Harter, “Parameterisation of a stochastic model for human face identification,” in *Proc. 2nd IEEE Workshop on Applications of Computer Vision*, Sarasota FL, Dec. 1994.

[17] A.M. Martinez and R. Benavente, “The ar face database,” *CVC Tech. Report # 24*, Jun. 1998.