Combining Face and Iris Biometrics for Identity Verification

Yunhong Wang[1], Tieniu Tan[1], and Anil K. Jain[2]

[1] National Lab of Pattern Recognition, Institute of Automation,Chinese Academy of Sciences
P.O. Box 2728, Beijing 100080, P.R. China
{wangyh,tnt}@nlpr.ia.ac.cn
[2] Department of Computer Sciences Engineering, Michigan State University
East Lansing, MI 48824
{jain}@cse.msu.edu

Abstract. Face and iris identification have been employed in various biometric applications. Besides improving verification performance, the fusion of these two biometrics has several other advantages. We use two different strategies for fusing iris and face classifiers. The first strategy is to compute either an unweighted or weighted sum and to compare the result to a threshold. The second strategy is to treat the matching distances of face and iris classifiers as a two-dimensional feature vector and to use a classifier such as Fisher's discriminant analysis and a neural network with radial basis function (RBFNN) to classify the vector as being genuine or an impostor. We compare the results of the combined classifier with the results of the individual face and iris classifiers.

1 Introduction

With increased need for reliable authentication schemes, the use of automatic identity verification systems based on biometrics has become widespread. Several airports are now equipped with biometric products. Face and iris recognition systems are among the top choices; because face recognition is friendly and non-invasive whereas iris recognition is one of the most accurate biometrics [7][11]. However, there are a number of practical issues that still need to be solved with both systems. The accuracy of face recognition is affected by illumination, pose and facial expression [15]. In many applications, face identification systems must be robust to these variations. In the case of iris recognition, the user must be cooperative. Further, iris images must meet stringent quality criteria, so the images of poor quality (e.g., iris with large pupil, or off center images) are rejected at the time of acquisition. Consequently, several attempts may be necessary to acquire the iris image, which not only delays the enrollment and verification, but also annoys the user. The rate of rejection of poor quality images is termed as the failure to enroll rate (FTE). Like any other biometric, the iris can change (e.g., as a result of eye disease), in which case, even a very good iris based identification system can fail.

Some of the above problems can only be solved, or at least their impact reduced, by fusing several biometric identification systems, such as face and iris recognizers. In general, by fusing several classifiers, the overall error rate (the false accept rate and the false reject rate) is known to go down [6]. This also reduces spoof attacks on the biometric system. The population coverage of a combined system is, in general, larger than the coverage of a standalone biometric, regardless of the accuracy of the latter; people with various disabilities may only be able to provide certain biometrics and not others. Thus, combining classifiers increases the number of people that can use the system. While it is true that a combined classifier requires the user to provide several biometrics during the acquisition stage, the combination of face and iris allows for simultaneous acquisition of face and iris images. Thus, in this particular case, no additional inconvenience is introduced. Finally, the use of the face recognizer in addition to the iris classifier, may allow people with imperfect iris images to enroll, reducing the enrollment failure rate.

There has been a substantial amount of work done on the combination of multiple classifiers [8][13]. Most of such work focuses on fusing 'weak' classifiers, for the purpose of increasing the overall performance. However, the advantages outlined above warrant the combination of existing 'strong' classifiers, for purposes other than increased performance. Not much work has been done in this direction. In this paper we develop a fused face-iris verification system which overcomes a number of inherent difficulties of the standalone classifiers. We compare the results of the combined classifier with the results of the individual face and iris classifiers.

2 Face Verification and Iris Verification

2.1 Face Verification

Among various face recognition algorithms, appearance-based approaches are the most popular. Examples include PCA [14], ICA [3] and LDA [4]. Here we use the Eigenface method as face matcher [14]. Let the ith sample face image be represented as an N-dimensional vector X_i, $i = 1, 2, ..., n$. The scatter matrix of all the n samples can be computed as $S = \sum_i (X_i - \mu)(X_i - \mu)^T$, where μ is the mean vector. The principal directions of S are the eigenvectors corresponding to the M largest eigenvalues of S, $M \ll N$. For each image X, we obtain a feature vector Y by projecting X onto the subspace generated by the principal directions. Images are then compared by means of their corresponding feature vectors.

2.2 Iris Verification

Iris recognition is receiving increased attention due to its high reliability [5]. The human iris is an annular region between the black pupil and the white sclera. The texture of iris are unique to each subject [5][10]. Daugman uses texture of iris image as features for classification[5]. The iris is first localized with two circles in the image. Then the iris part is unwrapped to a rectangular region where the iris texture is analyzed. The iris recognition system employed in this paper is based on an efficient algorithm that characterizes the critical points of local variations. The whole procedure of feature extraction includes two steps. First, a set of one-dimensional (1-D) intensity signals is constructed to effectively characterize the most important information in the original two-dimensional (2-D) image. Second, using a particular class of dyadic wavelets, a position sequence of local sharp variation points in such signals is recorded. For more details of the iris recognition system, see [9].

3 Fusion Approaches

The verification problem using a combination of classifiers (fused) can be formulated as follows. Classify a test sample S into one of the following two classes: ω_0 (genuine) or ω_1(impostor). If x_1 and x_2 are the outputs (matching distances) of the individual classifiers, then assign $S \rightarrow \omega_j$ if

$$P(\omega_j \mid x_1, x_2) = \max_{k=0}^{1} P(\omega_k \mid x_1, x_2),\, j = 0, 1 \tag{1}$$

where $P(\omega_j \mid x_1, x_2)$ denotes the posteriori probability of ω_k given x_1 and x_2. There are two different strategies that we employ for fusing the classifiers. One strategy is to compute either an unweighted or weighted sum of x_1 and x_2, and to compare the result to a threshold. The second strategy is to treat x_1 and x_2 as a two-dimensional feature vector. We then use a classifier such as Fisher's discriminant analysis and a radial basis function neural network (RBFNN) to decide whether the vector (x_1, x_2) represents a genuine sample or an impostor. We present the details of these methods below.

3.1 Weighted Sum Rule

The performances of different classifiers are different, so it is necessary to use different weights to combine the individual classifiers. Here we examine the performance (FAR and FRR) of each classifier to compute the weights. The FAR and FRR of the matchers are different at different thresholds. We sum the results of individual matchers with different weights at different thresholds. The weights are set according to the performance of individual matchers. Smaller error rates result in larger weights. The weights are computed by:

$$W_u = \frac{1 - (FAR_u + FRR_u)}{2 - (FAR_v + FRR_v + FAR_u + FRR_u)} \tag{2}$$

where $u = 1, 2$, $v = 1, 2$ and $u \neq v$. Note that $W_u + W_v = 1$.

The Weighted sum rule is defined as: assign $S \rightarrow \omega_j$ if

$$\sum_{i=1}^{2} W_i P(\omega_j \mid x_1, x_2) = \max_{k=0}^{1} \sum_{i=1}^{2} W_i P(\omega_k \mid x_1, x_2)\ . \tag{3}$$

3.2 Fisher Discriminant Analysis

If we treat the face and iris matcher outputs x_1 and x_2 as a feature vector $X = (x_1, x_2)$, then we can use any of the known classifiers to determine the separation boundary between the impostor and genuine samples. If we seek a linear boundary, then it is known that the line that, under the Gaussian assumption, best separates the two classes can be computed as $W^* = S^{-1}(m_1 - m_2)$, where m_1, m_2 and S are defined as

$$S = \sum_{x \in D_0} (x - m_1)(x - m_1)^T + \sum_{x \in D_1} (x - m_2)(x - m_2)^T \tag{4}$$

Here $m_1 = \frac{1}{n}\sum_{x \in D_0} x_1 \quad , m_2 = \frac{1}{n}\sum_{x \in D_1} x_2$, and D_0 and D_1 denote the set of samples from the genuine and impostor classes, respectively. The Fisher based fusion method can be expressed by the following equation:

$$x \in \omega_1 \ , \ if \ \ Y \geq Y_0, \ \ else \ \ x \in \omega_0 \ . \tag{5}$$

3.3 RBF Neural Network Based Fusion Method

Again, forming a vector (x_1, x_2) from the individual outputs of the face and iris classifiers, we use a RBF neural network for classification. We chose RBFNN over other types of multilayer-perceptron neural networks, because it had the best performance in our experiments.

We use two nodes in the input layer. One of the advantages of RBF neural network is that not only positive samples but also negative samples can be learned. For identity verification problem, the number of genuine samples is far fewer than the number of impostors. As the typicality of the subject is different, their genuine and impostor matching distances are different. It is necessary to design different RBF classifiers to fuse the individual matchers results. Since we are dealing with a verification problem, we can build an individual neural network for each subject in the database.

4 Experimental Results

4.1 Database

We have used two databases in our experiments, each containing face images and iris images. Aside from the number of subjects in each database, the main difference is that while the iris images in Database 1 are of very high quality, those in Database 2 are of somewhat lower quality.

Database 1. We collected the face images in Database 1 from the ORL [12], MIT [1], Yale [2] and NLPR databases. While the first three are well known public domain face databases, NLPR consists of face images taken in our lab at two different time instants. Examples of typical face images from the NLPR database are shown in Figure 1. The ORL database contains 40 subjects and 400 images. The MIT database contains 16 subjects and 432 images. The Yale database contains 15 subjects and 165 images. The NLPR database contains 19 subjects and 266 images. For each subject we selected 5 images, yielding a total of 450 face image for 90 subjects. The integrated database is composed of faces with reasonable variations in expression, pose and lighting.

There are no public domain iris image databases. The iris images used in our system come from the NLPR iris database. The database includes 2,096 iris images corresponding to 210 subjects, captured by an iris acquisition system developed at NLPR. There are at least 5 images for one eye. Iris images of left and right eyes are known to be different. Since not every individual provided iris images of both the eyes, there are 303 different classes from 210 subjects. The images were acquired during two different sessions, one month apart. Figure 2 shows some of the iris images from this database.

Obviously, the ORL, MIT and Yale face databases do not come with corresponding iris images, so to each face image, we assign an arbitrary (but fixed) iris class. Thus we obtain a database of 90 subjects, with 5 face images and 5 iris images per subject.

Fig. 1. Sample face images in the NLPR database.

Database 2. To illustrate the enrollment failure and its effect on the overall verification accuracy, we use some of the poor quality iris images that would normally be rejected in an operational iris verification system. We use 40 subjects and 400 iris images, 10 images per subject. The face database is the ORL database that includes 40 subjects and 400 images with 10 images for every subject.

(a) (b)

Fig. 2. Iris Database: (a) Iris images with good quality in DB1, (b) Iris images with poor quality in DB2.

4.2 Experimental Results

Results on DB1. The verification rate of the standalone iris classifier is very high. Most of the FAR and FRR rates at various thresholds are zero. It is difficult to show and compare the FAR and FRR rates using the ROC curve. Therefore, we simply calculated the total error rate (i.e., FAR+FRR). The outputs of RBF neural network are near 0 or 1, so it is not appropriate to express it as FAR and FRR at every threshold. The results based on the total error rate of the various verification systems are shown in Table 1. It is apparent from the distribution of matching distances that fusion improves the separation of genuine scores from the impostor scores. As the ROC curves can not be compared, we analyze the range of threshold values that result in high accuracy. The range of threshold values that can be selected for high verification rates is larger for the fused matcher compared to the individual matchers. This result can also be found from the distribution of distances in Figure 3.

There are a number of threshold values that result in a zero error, for iris and fused systems. The advantage of fusion is that we can get a larger range of operating points

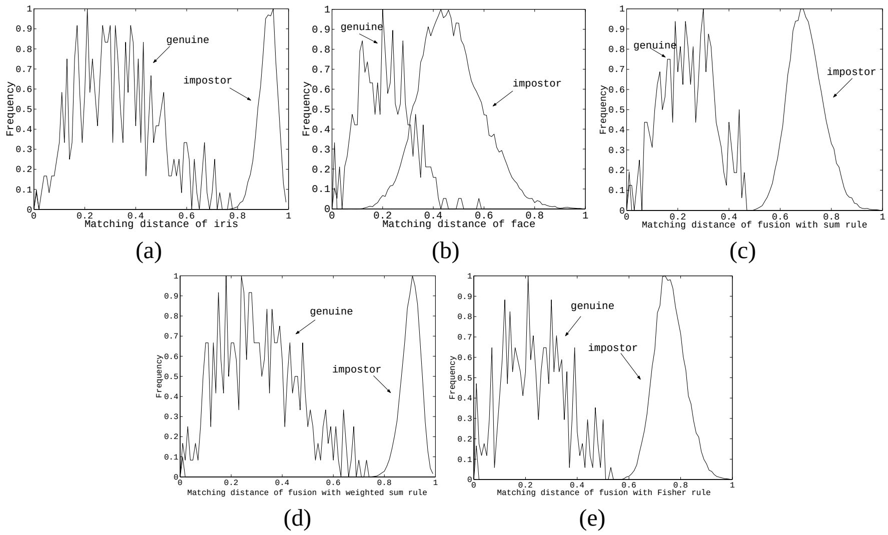

Fig. 3. Distribution of matching distances of (a) iris, (b) face, (c) the sum rule, (d) the weighted sum rule, (e) the Fisher rule.

(thresholds) with zero errors. Fusion systems are robust in this respect. The weighted sum rule has the best performance, followed by the Fisher discriminant. For the RBF neural network based fusion, we train a neural network for every subject. The total error rate is 0 (threshold is from 0.2 to 0.8).The neural network based approach obtains the highest verification accuracy here.

Table 1. Total Error rates of face, iris and fused systems

Threshold	Face	Iris	Sum Rule	Weighted Sum Rule	Fisher Rule
0.15	0.0245	0.0202	0.0244	0.0244	0.0080
0.25	0.0262	0.0138	0.0213	0.0213	0.0049
0.35	0.0298	0.0092	0.0106	0.0106	0.0019
0.45	0.3373	0.0048	0.0019	0.0031	5.2469e-4
0.55	0.9466	0.0013	0	0	0
0.65	0.9587	4.687e-4	0.0669	0	0.0328
0.75	0.9647	0	0.9550	0	0.4928
0.85	0.9681	0.0531	0.9669	0.0981	0.9280
0.95	0.9714	0.8300	0.9738	0.9738	0.9875

Results on DB2 The performance of iris verification system and fusion system is not as high on DB2 as it is for DB1 because some poor quality iris images are included in DB2. Now we can compare the matchers by means of the ROC curves shown in Figure 4. Based on Figure 4, the weighted sum rule has the highest verification accuracy. The performance of the Fisher classifier is similar to the performance of the weighted sum rule. The sum rule performed worse than the weighted sum rule and the Fisher classifier. The error rate of the RBF neural network classifier is 0.0024. Compared to the minimum error rates of the various verification systems, the RBF fusion method has the highest verification accuracy.

Although the iris system has high accuracy, it is not perfect. Figure 2 (b) shows some samples on which the standalone iris classifier failed, but they were correctly recognized by the RBFNN fused system. There are a number of similar situations in DB2. The reason these images were correctly classified by a combined classifier was that the classification does not depend entirely on a single modality. An operational iris verification system would reject poor quality images like the ones in Figures 2 (b), and require the user to try again to get a better iris image. Our system demonstrates that fusion is a way to decrease the enrollment failure rate and reduce verification time. At the same time, a large number of subjects that were misclassified by the standalone face verification system were correctly identified by the fusion system.

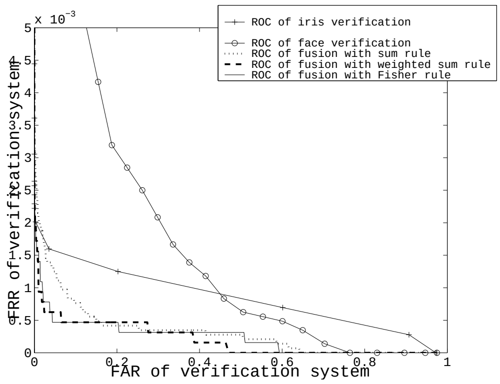

Fig. 4. ROC curves of verification systems

5 Conclusions

We have designed and built an identity verification system based on the fusion of face and iris data. The significance of fusing these two biometrics is more than the improvement in verification accuracy. Enlarging user population coverage and reducing enrollment failure are additional reasons for combining face and iris for verification. We have used two strategies for fusion: (i) weighted/unweighted summation of the outputs, x_1, and x_2 of the standalone classifiers, and (ii) treating (x_1, x_2) as a 2D feature vector, and using a Fisher discriminant analysis classifier and a neural network classifier. Fusion based on the RBF neural network produced the highest verification accuracy. Weighted sum rule is the best approach when compared with the sum rule and the Fisher rule. In Database 1, the performance of the fusion is worse than the performance of the iris standalone system except for the weighted sum rule. Database 2 includes some poor quality iris images, which will be rejected by an operational iris verification system. These images are accepted and correctly classified by the fused classifier, which means that the enrollment failure rate can be decreased by fusion, while maintaining a high accuracy.

Acknowledgements

Thanks to Silviu Minut for a careful reading of this paper, to Li Ma and Yuchng Fang for providing iris and face recognition algorithms, and to Arun Ross for his suggestions regarding fusion approaches. This research was supported by a grant from the NSF IUC on Biometrics (CITeR), at West Virginia University, the Chinese National Hi-Tech R-D Program (Grant No. 2001AA114180), and the NSFC (No.69825105).

References

1. MIT Face Database, ftp://whitechapel.media.mit.edu/pub/images.
2. Yale University Face Database, http://cvc.yale.edu/projects/yalefaces.html.
3. M. S. Bartlett, H. M. Lades, and T. J. Sejnowski. Indepent component representations for face recognition. In *Proc. of SPIE*, volume 3299, pages 528–539.
4. P. N. Belhumer, J. P. Hespanha, and D. J. Kriegman. Eigenfaces vs. fisherfaces: Recognition using class specific linear projection. *IEEE Trans. on PAMI*, 19(7):711–720, 1997.
5. J. G. Daugman. High confidence visual recognition of persons by a test of statistical independence. *IEEE Trans. on PAMI*, 15(11):1148–1161, 1993.
6. L. Hong and A.K. Jain. Integrating faces and fingerprints for personal identificatio. *IEEE Trans. on PAMI*, 20(12):1295–1307, 1998.
7. A.K. Jain, R.M. Bolle, and S. Pankanti. *Biometrics: Personal Identification in a Networked Society*. Kluwer, 1999.
8. Josef Kittler, Mohamad Hatef, Robert P. W. Duin, and Jiri Mates. On combining classifiers. *IEEE Trans. on PAMI*, 20(3):226–239, 1998.
9. Li Ma, Tieniu Tan, and Yunhong Wang. Efficient iris recognition by characterizing key local variations. submitted to IEEE Trans. on Image Processing.
10. Li Ma, Yunhong Wang, and Tieniu Tan. Iris recognition using circular symmetric filters. In *Proc. International Conference on Pattern recognition*, volume 2, pages 414–417, 2002.
11. T. Mansfield, G. Kelly, D. Chandler, and J. Kan. Biometric product testing final report. Technical report, National Physical Laboratory of UK, 2001.
12. Ferdinando Samaria and Andy Harter. Parameterization of a stochastic model for human face identification. In *Proc. 2nd IEEE Workshop on Applications of Computer Vision*, 1994.
13. M. Skurichina and R. P. W. Duin. Bagging, boosting and the random subspace method for linear classifiers. *Pattern Analysis and Application*, 5(1):121–135, 2002.
14. M. Turk and A. Pentland. Eigenfaces for recognition. *Journal of Cognitive Neuroscience*, 1(1):71–86, 1991.
15. W. Zhao, R. Chellappa, A. Rosefeld, and P.J. Phillips. Face recognition: A literature survey. Technical report, Computer Vision Lab, University of Maryland, 2000.