Performance Evaluation of Face Verification: A Comparative Study on Different Classifiers

SHAHRIN AZUAN NAZEER[1], MAZUKI KHALID[2], NAZARUDDIN OMAR[3] and MAT KAMIL AWANG[4]
[1,3,4]TM Research & Development Sdn. Bhd,UPM-MTC, UPM, Serdang
[2]Universiti Teknologi Malaysia, City Campus, Kuala Lumpur
MALAYSIA
{[1]shahrin,[3]nazar,[4]kamil}@tmrnd.com.my, [2]marz@citycampus.utm.my

Abstract - *The task of face verification is to verify the identity or decide whether the a priori user is an impostor or not from the known a priori identity of the user. The paper presents the performance evaluation carried out using different classifiers for face verification. The paper initially describes the approaches used for the face representation, and classification of face verification system. It then evaluates the performance of the system by applying three types of classifier: template-based matching, artificial neural network classifier, and Bayesian classifier based on AT & T and local face datasets. The measures used for performance evaluation are the false acceptance rate (FAR) and false rejection rate (FAR). Based on the experimental results, the artificial neural network classifier provides promising results for face verification with FAR of 4.44% and FRR 4.50% using AT&T face dataset, and FAR of 3.88 and FRR 4.00 % using local face dataset.*

Keywords: Face Verification, Euclidean Distance, Normalized Correlation, Artificial Neural Network, Bayesian Classifier.

1 Introduction

The demand for reliable personal identification in computerized access control has resulted with an increased interest in biometrics to replace password and identification (ID) card. The password and ID card can be easily breached since the password can be divulged to an unauthorized user, and the ID card can be stolen by an impostor. Thus, the emergence of biometrics has addressed the problems that plague the traditional verification methods. Biometric which make use of human features such as iris, retina, face, fingerprint, signature dynamics, and speech can be used to verify a person's identity. The biometrics data have an edge over traditional security methods since they cannot be easily stolen or shared. The necessity for personal verification or authentication in the fields of private and secure systems made face verification as one of the main fields among other biometric technologies. The importance of face verification rises from the fact that a face verification system does not require the cooperation of the individual, while the other systems need such cooperation or being a passive, non-intrusive system for verifying personal identity. The task of face verification is to verify the identity or decide whether the a priori user is an impostor or not from the known *a priori* identity of the user. The face verification system proposed is based on the popular feature extraction method, Principal Component Analysis (PCA) which is used for face recognition in [1], [2], [3], [4], [5], [6], [7], [9]. Using PCA, the face images are projected or map into eigenpictures or eigenfaces. The eigenfaces which are extracted using PCA are then classified using three types of classifiers which are Template-based matching classifiers [14], artificial neural networks (ANN) classifier [10], [15], and Bayesian classifier [13].

The objective of this paper is to present the performance evaluation of face verification based on these classifiers. The remainder of this paper is organized as follows. Section 2 describes approaches used for preprocessing and classification of the proposed face verification system. Section 3 presents and discusses the experimental results and the conclusions are drawn in section 4.

2 Approaches

The proposed face verification system consists of preprocessing and classification modules.

2.1 Preprocessing

Preprocessing is used to normalize and enhance the face image to improve the recognition performance. It involves the normalization and localization of the input face images. Initially, the input face images are normalized to grayscale face images. Next, the face images are localized using affine transformation and cropped using fixed proportion to obtain face images of size 50 by 50 pixels. Since the cropped face images are represented as multi-dimensional data, PCA is used to reduce the dimensionality by extracting the principal components of the face image. The first principal component is the linear combination of the original dimensions that has the maximum variance; the *n*-th principal component is the linear combination with the highest variance, subject to been orthogonal to the n - 1 first principal components.

In mathematical terms, the principal components of the distribution of faces are the eigenvectors of the covariance matrix of the set of face images, which treat an image as a point or vector in a very high dimensional face space. Every eigenface in the database is presented as a vector of weights, which are obtained by projecting the image into eigenface components by a simple inner product operation. A test image, whose verification is required, is also represented by its vector of weights.

2.2 Classification

The purpose of the classification sub-module is to map the feature space of a test data to a discrete set of label data that serves as template. The classification techniques used are Template-based matching classifiers: Euclidean Distance and Normalized Correlation, Artificial Neural Network classifier, and Bayesian classifier.

2.2.1 Template-based matching classifiers

The template-based matching classifiers used are Euclidean Distance and Normalized Correlation [8]. The Euclidean distance is the nearest mean classifier which is commonly used for decision rule is denoted as:

$$d_E(x, w_k) = \sqrt{(x - w_k)^T (x - w_k)} \qquad (1)$$

where the claimed client is accepted if $d_E(x, w_k)$ is below the threshold τ_{Ek} and rejected otherwise.

The normalized correlation is a decision rule based on the correlation score denoted as:

$$d_C(x, w_k) = \frac{|x^T w_k|}{|x||w_k|} \qquad (2)$$

where the claimed identity is accepted if $d_C(x, w_k)$ exceeds the threshold τ_{Ck}.

2.2.2 Artificial Neural Network (ANN) classifier

The ANN paradigm used is based on Multi-layer Perceptron (MLP) neural network. The MLP used is a form of non-linear network consisting of a set of inputs which forms the input layer, followed by one hidden layers of non-linear neurons and an output layer of non-linear neurons.

The MLP takes the features vector as input, and trains the network to learn a complex mapping for classification using error-correction back-propagation learning algorithm [11]. The learning algorithm involves repeatedly presenting the network with samples from a training set and adjusting the neural weights in order to achieve the required output. It is essentially a gradient descent method, where when adjusting the weight matrices, the direction is moved to the greatest descent. The training algorithm includes the momentum rate, used to speed up the convergence through any small local minima, and adaptive learning rate, used to prevent overshoot by adjusting the learning rate dynamically, usually starting with a large value and then decreases as it approaches the solution.

2.2.3 Bayesian classifier

Bayesian classifier is a technique that expands the eigenfaces method based on simple subspace-restricted norms using a probabilistic measure of similarity [12], [13]. The proposed similarity measure is based on a standard Bayesian analysis of image differences of two categories, which are intrapersonal variations, Ω_I, in the appearance of the same individual due to different expressions or lighting, and extra-personal variations, Ω_E, in appearance due to difference in identity.

The high dimensional probability density functions (pdf) for each class are then obtained from the training data using an eigenspace density estimation technique, and are subsequently used to compute a similarity measure based on the posteriori probability of membership in the intra-personal class. The probabilistic similarity measure, $S(I_1, I_2)$, between a pair of images, I_1, I_2 is denoted as:

$$S(I_1, I_2) = P(\Omega_I \mid \Delta) = \frac{P(\Delta \mid \Omega_I)P(\Omega_I)}{P(\Delta \mid \Omega_I)P(\Omega_I) + P(\Delta \mid \Omega_E)P(\Omega_E)} \qquad (3)$$

where $P(\Omega_I \mid \Delta)$ is the a posteriori probability given by Bayes rule, $P(\Delta \mid \Omega_I)$ and $P(\Delta \mid \Omega_E)$ are the estimates of the likelihood functions, and $P(\Omega)$ is the a priori knowledge regarding the two images being matched.

The problem is then solved using the maximum a posteriori (MAP) rule where two images are determined to be similar if $P(\Omega_I \mid \Delta) > P(\Omega_E \mid \Delta)$, or if $S(I_1, I_2) > 1/2$.

3 Experimental Results

The experiments are conducted to evaluate the performance of the face verification based on template-based matching classifiers using Euclidean Distance (ED) and Normalized Correlation (NC), Artificial Neural Network (ANN) classifier, and Bayesian classifier. The performance evaluation is conducted using the false acceptance rate (FAR) and false rejection rate (FAR). FAR is the case when an impostor, claiming the identity of a client, is accepted, whilst FRR is the case when a client claiming his true identity is rejected. The FAR and FRR are given by:

$$FAR = IA/I, \quad FRR = CR/C \qquad (4)$$

where IA is the number of impostor accepted, I is the number of impostor's trials, CR is the number of client rejected and C is the number of client's trials.

The experiments are conducted using two (2) datasets, which are AT&T face dataset, and local face dataset. The AT&T face dataset contains 10 different images of 40 distinct subjects. The images are grayscale with a resolution of 92 by 112 pixels. For the experiment, the images are rescaled to a resolution of 50 by 50 pixels, and divided into two parts: 5 images per subject for training, and another 5 images for testing with a total of 200 training images, and 200 testing images.

The local face dataset consists of 360 well-aligned frontal view face images. There are 20 different images of 18 distinct subjects. Similar to the experiments conducted using AT&T dataset, the face images for the local face dataset are rescaled to a resolution of 50 by 50 pixels and are divided into two parts: 5 images per subject are used for training, and another 15 images per subject are used for testing with a total of 90 training images, and 270 testing images.

The experimental results which are based on FAR and FRR are tabulated in Table 1. In Table 1, based on AT&T face dataset, the classifier using ANN provides the highest result with FAR of 4.44% and FRR of 4.50%, and using local face dataset, the ANN classifier also gives the highest result with FAR 3.88% and FRR of 4.00%. Thus, from the experimental results, the combination of both, PCA and ANN classifier gives the best result as compared to that of the other classifiers for both AT&T and local face datasets.

Dataset	Classifier	Threshold	FAR = FRR (%)	
			FAR	FRR
AT&T	ED	1.06	7.10	7.50
	NC	0.37	11.41	12.0
	ANN	0.21	4.44	4.50
	Bayesian	1.21	18.64	19.0
Local	ED	0.90	11.87	10.67
	NC	0.66	19.19	17.67
	ANN	0.60	3.88	4.00
	Bayesian	0.64	6.42	9.33

Table 1: Face verification results

4 Conclusions

The paper has presented a comparative study on three different types of classifier for face verification which are template-based matching classifiers using Euclidean distance and normalized correlation, ANN classifier, and Bayesian classifier. From the experimental results, it is shown that ANN classifier provides the highest overall results with FAR of 3.88%, and FRR of 4.00% using local face dataset, and FAR of 4.44% and FRR of 4.50% using AT&T dataset. Thus, in this paper we can conclude that the combination of both, PCA and ANN as a classifier is much superior compared to that of the template-based matching classifiers using Euclidean distance and normalized correlation, and Bayesian classifier for the overall performance for face verification.

References

[1] Kirby, M. and Sirovich, L., "Application of the Karhunen-Loeve Procedure for the Characterization of Human Faces", *IEEE Trans. On Pattern Analysis and Machine Intelligence,* Vol. 12, 1990.

[2] Turk, M. and Pentland, A., "Eigenfaces for Recognition", *Journal of Cognitive Neuroscience,* Vol. 3, pp.72-86, 1991.

[3] Stefano Arca, Paola Campadelli, Elena Casiraghi, Raaella Lanzarotti, "An Automatic Feature Based Face Authentication System", Proc. 16th Italian Workshop on Neural Nets (WIRN), pp. 120-126, 2005.

[4] Kyungim Baek, Bruce A. Draper, J. Ross Beveridge, Kai She, "PCA vs. ICA: A Comparison on the FERET Data Set", Proc. of the 6th Joint Conference on Information Science (JCIS) , pp. 824-827, 2002

[5] L. S. Balasuriya, N. D. Kodikara, "Frontal View Human Face Detection and Recognition", Proc. of the International Information Technology Conference (IITC), 2001.

[6] Bruce A. Draper, Kyungim Baek, Marian Stewart Bartlett, J. Ross Beveridge, "Recognizing faces with PCA and ICA." *Computer Vision and Image Understanding,* vol. 91(1-2), pp.115-137, 2003.

[7] P. J. B. Hancock, V. Bruce and A. M. Burton, "Testing Principal Component Representations for Faces", Proc. 4th Neural Computation and Psychology Workshop, 1997.

[8] Kilian Q. Weinberger, John Blitzer and Lawrence K. Saul, "Distance Metric Learning for Large Margin Nearest Neighbor Classification", *Neural Information Processing Systems (NIPS),* 2005.

[9] Wendy S Yambor, Bruce A. Draper J. Ross Beveridge, "Analyzing PCA-based Face Recognition Algorithms: Eigenvector Selection and Distance Measures", Proc. 2nd Workshop on Empirical Evaluation in Computer Vision, 2000.

[10] S. Lawrence, C. L. Giles, A. Tsoi, and A. Back, "Face recognition: A convolutional neural-network approach," *IEEE Trans. on Neural Networks,* vol. 8, pp. 98-113, January 1997.

[11]Fausett, Laurene V., *Fundamentals of neural networks: architecture, algorithms, and application,* Prentice-Hall, Inc., 1994.

[12]Bishop, C. M., "Bayesian PCA". In M. S. Kearns, S. A. Solla, and D. A. Cohn (Eds.), *Advances in Neural Information Processing Systems,* Vol. 11, pp.382–388. MIT Press, 1999.

[13] Moghaddam, B. Wahid, W. Pentland, A., "Beyond eigenfaces: probabilistic matching for face recognition", Proceedings. Third IEEE International Conference on_Automatic Face and Gesture Recognition, pp. 30-35, 14-16 Apr 1998.

[14] Brunelli, R., and Poggio, T. "Face Recognition: Features versus Templates", IEEE Transactions, PAMI, 15(10). pp. 1042-1052, 1993.

[15] Temdee, P., Khawparisuth, D., and Chamnongthai, K., "Face Recognition by Using Fractal Encoding and Backpropagation Neural Network", ISSPA '99, Brisbane, Australia, pp. 159-161, August 1999.