# FACE VERIFICATION WITHOUT FALSE ACCEPTANCE

Rosmawati Nordin
*Fakulti Teknologi Maklumat Dan Sains Kuantitatif, UiTM, Shah Alam, 40000 Selangor*
*roswati@tmsk.uitm.edu.my*

Md Jan Nordin
*Fakulti TeknologiDan Sains Maklumat, UKM, Bangi, 43600 Selangor*
*jan@ftsm.ukm.ny*

**Abstract.** Principal Components Analysis (PCA) and Linear Discriminant Analysis (LDA) are two popular approaches in face recognition and verification. The methods are classified under appearance-based approach and are considered to be highly-correlated. The last factor deems a fusion of both methods to be unfavorable. Nevertheless the authors will demonstrate a verification performance in which the fusion of both method produces an improved rate compared to individual performance. Tests are carried out on FERET (Facial Recognition Technology) database using a modified protocol. A major drawback in applying LDA is that it requires a large set of individual face images sample to extract the intra-class variations. In real life application data enrolment incurs costs such as human time and hardware setup. Tests are therefore conducted using virtual images and its performance and behaviour recorded as an option for multiple sample. The FERET database is chosen because it is widely used by researchers and published results are available for comparisons. Performance is presented as the rate of verification when false acceptance rate is zero, in other words, no impostors allowed. Initial results using fusion of two verification experts shows that a fusion of T-Zone LDA with Gabor LDA of whole face produces the best verification rate of 98.2% which is over 2% improvement compared with the best individual expert.

## 1. Introduction

The absolute objective of a pattern recognition system is to achieve the best classification performance. Activities involved includes devising a classification scheme that best suited the given problem. In order to find the best classifier experiments are normally carried out using the data fed into the various classification techniques and comparing the outcome. The classiffier that produces the best results in terms of recognition or verification rate will be chosen. Nevertheless one observation is that even though one classifier is considered the best, the pattern of missclassification is not necessarily similar or overlap indicating that the decisions from the different classifier can be used to further increase the performance of the chosen classifier.

Another idea that resulted from the observation is that to fuse or combine decisions from different classifiers to achieve the final decision. Various classifier combination scheme has been proposed and tested whereby the general consensus is that the performance of classifier fusion is consistently better compared to a single classifier (Kittler et al 1997, Jain et al 2000 and Duin 2002). Nevertheles it is still unclear why a certain combination performs better than the other.

## 2. PCA and LDA For Face Verification

Principal component Analysis (PCA) is a classical technique for multivariate analysis. It was adopted for use in face detection and recognition by Turk and Pentland (1991) where faces are represented in a low-dimensional feature space known as eigenspace. An eigenspace is computed by estimating the eigenvectors of the covariance matrix of a training face set. The eigenvectors corresponding to the largest eigenvalues are taken as the principal components of the eigenspace and capture the main modes of global variation in the training data set. These principal components are also known as eigenfaces. Computationally, PCA yields a matrix $\mathbf{U} = [\mathbf{u}_1 \ \mathbf{u}_2 \ldots \mathbf{u}_k]$ whose columns are the $K$ most significant eigenvectors of the training data's covariance matrix, corresponding to the $K$ largest eigenvalues ($\lambda_1$, $\lambda_2$ …, $\lambda_k$). A novel face image, **x,** is represented by the linear projections onto the K-dimensional eigenspace given by a coefficient vector

$$\boldsymbol{\alpha} = \mathbf{U}^T(\mathbf{x} - \boldsymbol{\mu}) \qquad (1)$$

where $\boldsymbol{\mu}$ is the mean image of the training set. Typically, $K$ is between 30 and 50 which is one or two orders of magnitude lower than the dimensionality of the image space

Linear Discrimanat Analysis (LDA) also known as Fisher Discriminant Analysis searches for a linear transformation $W$ dari $R^m$ to $R^{m_{lds}}$ di mana $m_{lda} \leq q - 1$ ( $q$ is classes in training data) that maximises class separability J. There are several methyods to define separability of Linear transformation $W$ and the method widely used is

$$J(W) = \frac{\det\left|W^T S_b W\right|}{\det\left|W^T S_w W\right|}, \qquad (2)$$

where $S_b$ is the between class variation while $S_w$ is the within class variation.

## 3. Classifier Combination

Face verification is considered as two-class problem and normally solved based on template matching method. A score or distance value is measured between a query image and the template and compared to the threshold. The query image is accepted if the score is bigger (or smaller) than the threshold value.

Various level of classifier combination can be applied (Czyz et al 2003) : data level, feature extraction level or decision level. Experiments carried out are decision level classifier fusion using the basic classifers explained in section 2 for frontal images in grayscale values as well as gabor filtered images. Combinations methods tested are fixed rules sum, prodect and median. Detailed explanation can be found from Kittler (1998) and Kuncheva (2002).

The two main reasons for combining classifiers are efficiency and accuracy. An interesting issue in the research concerning classifier ensembles is the way they are combined. If only labels are available a majority vote (Kimura and Shridhar 1991) is used. If continuous outputs like posteriori probabilities are supplied, an average or some other linear combination have been suggested (Kittler et al 1997). It depends on the nature of the input classifiers and the feature space whether this can be theoretically justified. If the classifier outputs are interpreted as fuzzy membership values, belief values or evidence, fuzzy rules (Cho and Kim 1995), belief functions and Dempster-Shafer techniques (Xu et al 1992, Rogova 1994) are used. Finally it is possible to train the output classifier separately using the outputs of the input classifiers as new features (Wolpert 1992).

This study focusses on frontal face verification using multi-classifier approach at decision-level.

## 4. Face Database and Experiment Protocol

Facial images for the experiment are gathered from from FERET(2004) frontal grayscale images. The database stores a collection of face images in various pose and illumination. Images are grouped into different poses and the sets are used as queries in tests. The experiments carried out involve only the images in *fafb* probes and further partitioned as in figure 1. Figure 2 displays a few images from the database. Behaviours of the verifier when virtual images are added for the training images are also observed. The virtual images are derived from the translation and transformation of the original images as shown in figure 3. Virtual images provides options to increase the sample size of images in the database.

Wiskot et al (1997) introduced an image representation which is derived from Gabor jets convolusion in 5 frequencies and 8 orientations that is quite tolerant to variations in viewpoint. The same method is adopted as a comparison to PCA in grayscale representation. Figure 4 shows the convolusion result.

The fusion of classifiers involves combining six verification experts, one based on LDA (LDA-T), two based on PCA in grayscale images (PCA-T dan PCA-W), one based on PCA of Gabor features (GAB-W2) and two based on LDA of gabor features (GAB-W1 dan GAB-T).

| **Client(109)** | | | **Impostor(80)** | |
|---|---|---|---|---|
| **Training Set** | **Validation Set** | **Test Set** | **Validation Set** | **Test Set** |

**Figure 1 : Data Set Partitions**


**Figure 2 : Face Image Examples**

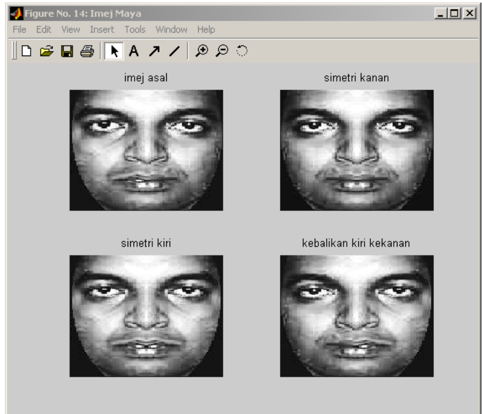


**Figure 3 : Vitual Image Examples**

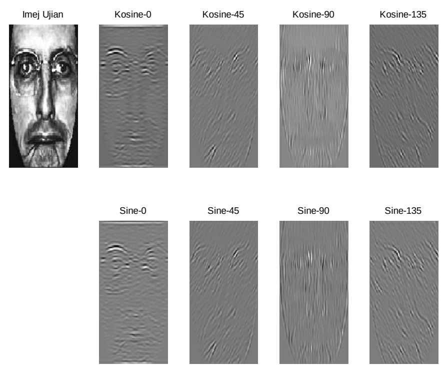


**Figure 4 : Gabor-convoluted Images**

In all the verification experts, the initial steps include localisation to determine facial area. The areas examined in the experiment is the whole face excluding hair and shoulder as well as the T-Zone area which is the eyes and nose region. These areas will be cropped and fed to the verifier. Various methods are available for example using the T-Zone template (Nordin and Nordin, 2003).

## 5. Experimental Results

In principle, the possible number of combinations follow the the formula $^{\mathbf{n}}\mathbf{C}_{\mathbf{r}}$ where in this experiment r is the number of experts which is 1 (6 experts individually) up to 2,3,4,5,6 (all experts combined). Table 1(a) shows the number of combinations. The result reported in this paper covers only fusion of two classifiers using sum combination rule.

**Table 1(a) : Number of Combinations**

| r | $^{n}Cr$ |
|---|---|
| 1 | 6 |
| 2 | 15 |
| 3 | 20 |
| 4 | 15 |
| 5 | 6 |
| 6 | 1 |

From table 1(b) it shows that PCA-T verifier gives the best verication rate for all training set size in which the best is 82% when there are four images for every client in the training set. Even though the increase in size is due to virtual images, verification experts based on PCA of grayscale images demonstrate a better performance. The verification performance is indicated by TAR (True Acceptance Rate) that is the percentage of Successful client acces over the number of clients.

**Table 1(b) Verification Performance at threshold value FAR=0**

| Classifier | N=1 | N=2 | N=3 | N=4 |
|---|---|---|---|---|
| PCA-T | **53.21** | **78.90** | **87.16** | **96.33** |
| LDA-T | 38.53 | 70.64 | 81.65 | **96.33** |
| GAB-T | 9.17 | 47.71 | 56.88 | 60.55 |
| PCA-W | 46.79 | **78.90** | 78.90 | 79.82 |
| GAB-W1 | 16.51 | 61.47 | 64.22 | 65.14 |
| GAB-W2 | 27.52 | 63.30 | 63.30 | 63.30 |

Nevertheles results from separate experiments show a significant improvement of LDA over PCA for Gabor convoluted images in the wholeface as well the T-Zone regions compared to grayscale images.

**Table 2 Performance of the fusion of verification experts using Sum rules at threshold FAR=0**

| Classifier | N=1 | N=2 | N=3 | N=4 |
|---|---|---|---|---|
| PCA-W, PCA-T | 56.88 | **84.40** | **92.66** | 97.25 |
| PCA-W, GAB-W2 | 53.21 | 82.57 | 82.57 | 82.57 |

| | | | | |
|---|---|---|---|---|
| PCA-W, LDA-T | 52.29 | 82.57 | 90.83 | 97.25 |
| PCA-W, GAB-W1 | 47.71 | 80.73 | 83.49 | 85.32 |
| PCA-W, GAB-T | 49.54 | 80.73 | 84.40 | 84.40 |
| PCA-T, GAB-W2 | **57.80** | **84.40** | **92.66** | 97.25 |
| PCA-T, LDA-T | 55.90 | 83.49 | 88.99 | 96.33 |
| PCA-T, GAB-W1 | 47.71 | 80.73. | 83.49 | 85.32 |
| PCA-T, GAB-T | 54.13 | 81.65 | 88.99 | 96.33 |
| GAB-W2, LDA-T | 46.79 | 81.65 | 88.99 | 97.25 |
| GAB-W2, GAB-W1 | 33.94 | 75.23 | 76.15 | 76.15 |
| GAB-W2, GAB-T | 29.36 | 70.64 | 75.23 | 79.82 |
| LDA-T, GAB-W1 | 42.20 | 77.98 | 87.16 | **98.17** |
| LDA-T, GAB-T | 40.37 | 72.48 | 85.32 | 96.33 |
| GAB-W1, GAB-T | 22.94 | 66.97 | 71.56 | 77.98 |

Results from table 2 indicate improve performance in verification for all fusion of experts with increased training size. For N=3 and N=4 which add virtual images in the training sets indicate that they played a positive role in the improvements. Another observation from the results of fusion of two verification experts is that the fusion of the two best individual verifiers (PCA-T and LDA-T) does not produce the best performance. The best performance is achieved by the fusions of LDA-T and GAB-W1.

## 6. Conclusions

This paper provides a brief explanation about face verification using fusion of verifiers at decision level. FRVT (2002) reported that the state of the art error rate should be in the range of 10% False Acceptance Rate(FAR) with 1% False Rejection Rate(FRR). The two errors are counter-complementary in the sense that reduction in one error rate increases the other. This paper reports the verification performance when FAR=0% which is quite impossible to achieve in the real application. No guidelines exist on how to choose verifiers or classifiers to be combined which leaves the tasks based on intuition or heuristics. This activities can be regarded as an art. Nevertheless the verifiers chosen in the experiment are listed as the top five performers in the various competitions conducted as reported by Phillips et al (1997) and Matas et al (2000).

## References

[1] S.B. Cho,. and J.H. Kim, “Multiple Network Fusion Usimg Fuzzy Logic”, *IEEE Trans.,Neural Networks* vol 6, 1995, pp. 497-501.

[2] J. Czyz, J.Kittler and L.Vandenthorpe,. "Multiple classifier combination for face-based identity verification", *Proc. of Int. Conf. on Pattern Recognition*, 2003.

[3] R. P. W Duin,. "The Combining Classifier: to train or not to train", *Proc. of Int. conf. on Pattern Recognition*, 2002.

[5] A.K. Jain, R.P.W Duin, and J. Mao, "Statistical Pattern Recognition: A review", IEEE Trans. On Pattern Recognition and Machine Intelligence, vol.22(1), 2000

[6] F. Kimura and M. Shridhar, "Handwritten Numerical Recognition Based on, Multiple Algorithms," *Pattern Recognition*, vol.24(10) 1991,pp. 969-983.

[7] J.Kittler, A. Hojjatoleslami and T. Windeatt,."Weighting Factors in Multiple Expert Fusion," *Proc. British Machine Vision Conf.*, Colchester, England, 1997, pp. 41-50.

[8] J. Kittler, "Combining Classifiers: A Theoretical Framework", *Pattern Analysis And Applications* 1 , 1998, pp 18-27.

[9] L. Kuncheva,. "A theoretical study on six fusion strategies", *IEEE Trans. On Pattern Recognition and Machine Intelligence*, 2002, vol.24(2), pp 281-286.

[10] J.Matas, M. Hamouz, K.Jonsson., J.Kittler, Y.Li, C. Kotroupolous, A.Tefas,., I. Pitas, T. Tan, H. Yan, B. Smeraldi, J. Bigun, N. Capdevielle, W. Gerstner, S. Ben-Yacoub, Y. Abduljaoued and E. Majoraz,. "Comparison of face verification results on the xm2vts database", *Proceeding of the International Conference on Pattern Recognition (ICPR)*, 2000.

[11] R. Nordin and M.J. Nordin, "Face Detection For Recognition From A Single Frontal Grayscale Image ", *Proc. of Int. Conf on ROVISP*, 2003, pp 135-139.

[12] P.J. Phillips, H. Moon, R. Rauss, and S. Rizvi, "The FERET evaluation methodology for face recoognition algorithm", *Proc. of Int. Conf. on CVPR*, 1997.

[13] G. Rogova, "Combining the Results of Several Neural Network Classifiers," *Neural Networks*, 1994, vol. 7(5), pp. 777-781.

[14] M. Turk and A. Pentland, 1991. "Eigenfaces for recognition",. *Journal of Cognitive Neuroscience*, 1991, vol.3, pp71-86.

[15] L. Wiskott, J.M.. Fellous, N. Kruger, C. Malsburg, "Face recognition by elastic bunch graph matching", *IEEE Trans. On PAMI*,1997, vol.19(7), pp 775-779.

[16] D.H. Wolpert,. "Stacked Generalization", *Neural Networks*, 1992 , vol. 5, pp 241-260.

[17] L. Xu, A. Krzyzak and C.Y. Suen, "Methods of Combining Multiple Classifiers and Their Applications to Handwriting Recognition," *IEEE Trans. Systems, Man, and Cybernetics*, 1992, vol. 22:3, pp. 418-435.

[18] The Facial Recognition Technology (FERET) Database. 2004. (online) http://www.itl.nist.gov/iad/humanid/feret/(10 Nov 2006).

[19] FRVT 2002:Face Recognition Vendor Test Sept. 2003 (online) http://www.frvt.org (15 June 2007).