Face Authentication Using Client Specific Fisherfaces

J Kittler

Centre for Vision Speech and Signal Processing
University of Surrey, Guildford, Surrey GU2 5XH, England

Telephone: +44(0)1483 879294 Fax: +44(0)1483 876031
Email: J Kittler@ee.surrey.ac.uk

November 1999

Abstract

In the paper we propose a one dimensional client specific fisher face representation for personal identity verification. This novel LDA approach contrasts with the conventional LDA representation which involves multiple shared fisher faces. The method provides two measures for authentication: a distance to the client template, and a distance to the mean of impostors. These two decision scores are combined to achieve significant performance gains.

The method is tested on the XM2VTS database according to the internationally agreed Lausanne protocol. The demonstrated performance superiority is not the only advantage of the proposed method. Additional features of practical significance include the simplicity of training, as for large user databases the proposed technique requires only a matrix multiplication of the client mean vector. Moreover, the client enrollment is insulated from the enrollment of other clients. This opens the possibility to use other than the centralised architecture for the personal identity verification system and in fact smart card processing becomes a reality without any need to restrict the representation framework and therefore the representational capacity of the system. Finally the speed of probe testing is more than two orders of magnitude faster than that achieved by conventional PCA and LDA methods as the proposed techniques involves only a single fisher face per client. These attractive properties make the method ideally suited for both representation and authentication in personal identity verification systems.

1 Introduction

The human face plays a major role in conveying personal identity. The development studies of the human face recognition abilities reported in [9] show that even a few minutes after the birth a baby is able to track face-like objects which suggests innate, hardwired capabilities. Although the face recognition abilities continuously develop during the childhood, the recognition performance for upright faces appears to improve sharply between the ages of 6 and 10. This improvement is not matched by the ability to discriminate upside down faces which indicates that most of the recognition ability is learned and adapted to a typical situation in which face recognition is likely to be performed.

This unique capacity of the human nervous system, coupled with the fact that there are many situations where facial photographs are the only information available about individuals, makes the face recognition modality the subject of extensive studies aimed at both understanding the process and being able to emulate it by a machine. The latter is of a particular interest as personal identity verification is the enabling technology for many applications such as i building access control, recognition of suspects from CCTV video against a gallery of faces drawn from police face databases, mug shots matching, credit card user verification, access to teleservices such as bank teller machines, etc.

The problem of computerised personal identity verification has received considerable attention over the past decades and a large number of approaches have been proposed in the literature. Though the early strategy for face identification was based on geometrical features such as nose width and length, mouth position and chin shape, etc, [6, 10, 5] nearly all of the applicable approaches developed in recent years are based on holistic representations known as templates. A comparison of geometrical feature-based matching with template (statistical feature) matching presented by [2] favours the statistical feature based matching approach. Surveys of statistical template matching techniques can be found in [3].

The most commonly used statistical representation for face recognition and verifica-

tion is the Karhunen-Loeve (KL) expansion [7] known also as the Principal Component Analysis (PCA). Its application to the face recognition problem has been pioneered by Sirovich and Kirby [31, 17] but the approach has been popularised by Turk and Pentland [34] where the PCA bases are referred to as eigenfaces. Since then the eigenface method has been widely used by many researchers. Examples can be found in [28, 30, 11]. A detailed analysis of PCA-based face recognition algorithms which share the same eigenface representation but differ in terms of the classification strategy can be found in [27].

Although PCA is very effective for information compression, it does not guarantee the most efficient compression of discriminatory information. More recently, the linear discriminant analysis (LDA) [35, 7] has been adapted to face recognition by Belhumeur et al. [1]. The LDA representation bases, referred to as "fisherfaces", were demonstrated to outperform the PCA representation. As LDA involves the eigenanalysis of a product of two matrices, one of which is inverted, Swets [32] advocated that the face data should first be projected into a PCA space to ensure that the matrix being inverted is not rank deficient. The LDA-based features derived in this lower dimensional space were shown to be the most discriminating features. Face recognition or verification using fisher faces was studied by many authors including [13, 14, 32], [22, 36], [21], etc. All these papers use the conventional LDA subspace spanned jointly by all the client classes.

Many other approaches have been proposed in the literature such as active appearance models [4, 12], robust correlation [23, 15], support vector machine (SVM) [29, 26]. An interesting development is also the dynamic link architecture matching [20] which applies conventional image matching techniques locally after applying nonrigid transformations to correct for local deformations of a face due to a changing expression or mouth movement. However, it is far from clear at this stage whether these techniques will offer a superior solution relative to LDA in terms of both their recognition/verification performance and computational complexity.

All of above approaches are based on a global representation of both the training samples and the probe in a subspace of the training data space, called feature space.

The distribution of the training data of the clients in this subspace induces a space tessellation [33] which defines the spatial extent of each class. The verification of a claimed identity is then tested by checking whether the probe image projects into an appropriate tessel of the representation space. The feature space tessellation is normally imposed by means of a global threshold which optimally separates each class from all the other classes. Alternatively, the feature space partitioning can be achieved using a client specific thresholding which can significantly improve the classification ability.

While client specific thresholding attempts to achieve a better adaptation to class specific distributions it is realised within the framework of a shared multiclass representation of all the face data. To our knowledge there is only one technique which departs from the common feature space approach and deploys client specific representations, namely the client specific SVM method of Jonsson et al [16]. Interestingly, even this approach builds the client specific bases in a common PCA feature space spanned by the training face image data.

In contrast to the common multidimensional representation framework of the standard PCA and LDA approaches, in this paper we propose one dimensional client specific fisher face representation. This novel LDA approach achieves a superior performance to any known method tested on the XM2VTS database according to the internationally agreed Lausanne protocol. In comparison with the next best approach which is based on client specific Support Vector Machine design, the performance gain of the proposed method is greater than two. Equally importantly, the performance superiority is not the only advantage of the proposed method. Additional features of practical significance are:

- the **simplicity of training**, which for large user databases requires only a matrix multiplication of the client mean vector;
- the **insulation of a client enrollment** from the enrollment of other clients. This opens the possibility to consider other than the centralised architecture for the personal identity verification system (i.e. architecture where client models are

stored and updated centrally) and

- **smart card processing** becomes a reality without the need to restrict the the representation framework and therefore the representational capacity of the system;

- the **speed of probe testing** being more than two orders of magnitude faster than conventional PCA and LDA methods as the proposed techniques involves only a single fisher face per client.

The paper is organised as follows. In Section 2 we develop the novel client specific LDA representation and discuss associated decision strategies. We argue that with the client specific fisher faces the problem of decision making can be formulated as either testing the hypothesis that the claimed identity is true or that the probe image belongs to an imposter. We also show that both tests can be applied simultaneously and the results fused to enhance the system performance. In Section 3 we list the benefits of the proposed method. Its performance is evaluated experimentally on a large data base of face images in Section 4. The data for experimentation is extracted from the public domain multimodal XM2VTS database designated for benchmarking algorithms for personal identity verification. The experiments are conducted using the Lausanne protocol. Finally the paper is drawn to conclusion in Section 5.

2 Theory

Let us consider the problem of designing a personal identity verification system for m clients. Suppose that for each client we have N_i samples of some biometric data for training. In here the focus will be on frontal face images but the methodology to be developed is applicable to any biometric modality. Accordingly, we shall assume that the data is appropriately registered and normalised photometrically. Note that in principle the number of training samples per client may differ. Thus the size of the training set will be $N = \sum_{i=1}^{m} N_i$.

The conventional approach to personal identity verification or to face recognition based on fisher faces is first to project the training data into the lower dimensional subspace spanned by the training images. This is required so that the within class scatter matrix that has to be inverted when searching for fisher faces is not rank deficient. The projection matrix can be found by means of principal component (PCA) analysis. Denoting the $j-th$ training image in a vector form as $\mathbf{z}_j$ the analysis commences by finding first of all the global mean $\mu = \sum_{j=1}^{N} \mathbf{z}_j$ and the mixture covariance matrix Σ as

$$\Sigma = \sum_{j=1}^{N} (\mathbf{z}_j - \mu)(\mathbf{z}_j - \mu)^T \tag{1}$$

If the dimensionality of the image vectors is larger than the number of training images N, the mixture covariance matrix Σ will have at most $n \leq N$ non zero eigenvalues. The eigenvectors of Σ associated with the nonzero eigenvalues define the subspace spanned by the training data and can be used for projecting the training images into a lower, n dimensional space fully spanned by the projected training data. This can be accomplished without any loss of information.

It should be noted that for computational reasons the practical procedure for finding the eigenvectors of matrix Σ differs from the above description so that the eigenvalue analysis can be solved in the $N \times N$ space rather than in the $D \times D$ space as the latter numerical analysis problem may not be feasible. However, the details of the feasible approach can be found elsewhere [35]. For the purpose of this paper we shall simply assume that the n eigenvectors $U = [\mathbf{u}_1, ..., \mathbf{u}_n]$ have been found and that the image data, after centralisation, has been projected into lower dimensional vectors $\mathbf{x}_j$ as

$$\mathbf{x}_j = U^T(\mathbf{z}_j - \mu) \;\; \forall j \tag{2}$$

Let us denote the mixture covariance matrix of the projected vectors by Φ, i.e.

$$\Phi = \frac{1}{N} \sum_{j=1}^{N} \mathbf{x}_j \mathbf{x}_j^T \tag{3}$$

The standard approach would now proceed first by defining the mean of each client

class $\omega_i,\ i = 1, .., m$ as

$$\nu_i = \frac{1}{N_i} \sum_{j=1}^{N_i} \mathbf{x}_j \quad \mathbf{x}_j \epsilon \omega_i \tag{4}$$

and the mean class scatter matrix S_B. The fisher faces $\mathbf{v}_k$ would then be obtained by finding the eigenvectors of matrix $\Phi^{-1} S_B$ associated with the non zero eigenvalues of this matrix product. There would be at most $m - 1$ such eigenvectors. Personal identity claims would then be tested by projecting a probe image $\mathbf{z}$ into the PCA space U, followed up by a projection onto fisher faces $V = [\mathbf{v}_1, ..., \mathbf{v}_{m-1}]$ as

$$\mathbf{y} = V^T U^T (\mathbf{z} - \mu) \tag{5}$$

The projected probe vector $\mathbf{y}$ would then be tested against the projected mean γ_i

$$\gamma_\mathbf{i} = V^T \nu_\mathbf{i} \tag{6}$$

of the claimed identity $\omega_\mathbf{i}$ using a suitable metric.

Note that in the conventional approach the representation space for the probe vector and client template is the same for all identity claims. In this paper we advocate a diametrically different approach by proposing the use of client specific representation spaces. This is achieved by formulating the personal identity verification problem as a two class problem: class ω_i is the $i - th$ client claimed identity and Ω is the imposter class (claim not accepted). It is interesting to note that the mean of the imposter class $\nu_\mathbf{\Omega}$ in the PCA space can be expressed as

$$\nu_\mathbf{\Omega} = \frac{1}{N - N_i} \sum_{j=1}^{N - N_i} \mathbf{x}_j \ \ \mathbf{x}_j \epsilon \omega_i \tag{7}$$

which can easily be shown to simplify to

$$\nu_\mathbf{\Omega} = -\frac{N_i}{N - N_i} \nu_i \tag{8}$$

Thus the mean of the imposter class is the global mean shifted in the opposite direction to that of the client mean. The magnitude of the shift is given by the ratio of the respective numbers of the client training face images and the total number of training faces.

This will normally be very small and the mean of the imposter class will always stay close to the origin irrespective of the claimed identity.

The between class scatter matrix M_i in this two class case will be given as

$$M_i = \frac{N - N_i}{N}(\frac{N_i}{N - N_i})^2 \nu_i \nu_i^T + \frac{N_i}{N}\nu_i \nu_i^T \tag{9}$$

which can be reduced to

$$M_i = \frac{N_i}{N - N_i}\nu_i \nu_i^T \tag{10}$$

The covariance matrix of the imposter class Φ_Ω estimated as

$$\Phi_\Omega = \frac{1}{N - N_i}\sum_{j=1}^{N-N_i}(\mathbf{x}_j + \frac{N_i}{N - N_i}\nu_i)(\mathbf{x}_j + \frac{N_i}{N - N_i}\nu_i)^T \quad \mathbf{x}_j \epsilon \omega_i \tag{11}$$

can be expressed in terms of matrix Φ by rewriting equation (11) as

$$\begin{aligned} \Phi_\Omega = \quad & \tfrac{1}{N-N_i}\{\textstyle\sum_{j=1}^{N}(\mathbf{x}_j + \tfrac{N_i}{N-N_i}\nu_i)(\mathbf{x}_j + \tfrac{N_i}{N-N_i}\nu_i)^T \\ & - \textstyle\sum_{j=1}^{N_i}(\mathbf{x}_j + \tfrac{N_i}{N-N_i}\nu_i)(\mathbf{x}_j + \tfrac{N_i}{N-N_i}\nu_i)^T\} \end{aligned} \tag{12}$$

where the vectors in the second sum belong to the client class. In fact the second sum is related to the covariance matrix Φ_i for the client class, i.e.

$$\Phi_i = \frac{1}{N_i}\sum_{j=1}^{N_i}(\mathbf{x}_j - \nu_i)(\mathbf{x}_j - \nu_i)^T \quad \mathbf{x_j} \epsilon \omega_i \tag{13}$$

as

$$\sum_{j=1}^{N_i}(\mathbf{x}_j + \frac{N_i}{N - N_i}\nu_i)(\mathbf{x}_j + \frac{N_i}{N - N_i}\nu_i)^T = N_i\left\{\Phi_i + \left(\frac{N}{N - N_i}\right)^2 \nu_i \nu_i^T\right\} \tag{14}$$

Simplifying (12) we finally find

$$\Phi_\Omega = \frac{1}{N - N_i}\left[N\Phi - N_i\Phi_i - \frac{NN_i}{N - N_i}\nu_i \nu_i^T\right] \tag{15}$$

The within class scatter matrix Σ_i is now obtained by weighted averaging the covariance matrices of the class of impostors and clients, i.e.

$$\Sigma_i = \frac{N - N_i}{N}\Phi_\Omega + \frac{N_i}{N}\Phi_i \tag{16}$$

Substituting from (15) and simplifying we finally obtain

$$\Sigma_i = \Phi - M_i \tag{17}$$

We are now in position to determine the linear discriminant for this two class problem by finding the eigenvectors of matrix $\Sigma_i^{-1} M_i$ associated with nonzero eigenvalues. In fact in the two class problem there is only one such vector $\mathbf{v}$ that satisfies

$$\Sigma_i^{-1} M_i \mathbf{v} - \lambda \mathbf{v} = 0 \tag{18}$$

with $\lambda \neq 0$ provided ν_i is non zero. As there is only one solution to the eigenvalue problem it can be easily shown that vector $\mathbf{v}$ can be found directly, without performing any eigenanalysis as

$$\mathbf{v} = \Sigma_i^{-1} \nu_i \tag{19}$$

This becomes apparent by substituting for $\mathbf{v}$ in (18) from (19) and for M_i from (10), i.e.

$$\frac{N_i}{N - N_i} \Sigma_i^{-1} \nu_i \nu_i^T \Sigma_i^{-1} \nu_i = \lambda \mathbf{v} \tag{20}$$

which also shows that the eigenvalue λ is given by

$$\lambda = \frac{N_i}{N - N_i} \nu_i^T \Sigma_i^{-1} \nu_i \tag{21}$$

Thus the overall client i specific linear discriminant transformation $\mathbf{a}_i$ is given as

$$\mathbf{a}_i = U \mathbf{v} \tag{22}$$

It defines the client specific fisher face for testing the claimed identity.

2.1 Decision strategy

It is now interesting to note that given a probe $\mathbf{z}$ the claimed identity ω_i can be tested in two different ways. First of all it can be tested against the client mean vector μ_i by centralising both by the global mean μ and projecting them on to the fisher face $\mathbf{a}_i$. The difference between these one dimensional projections will be indicative of the degree of

similarity (in fact dissimilarity) between the probe and the client model. In computing the difference, the effect of the centralisation of the probe and client mean vectors by the global mean is cancelled. Thus the test statistic (or score) to be used for decision making can simply be defined as

$$d_c = |\mathbf{a}_i \mathbf{z} - \mathbf{a}_i \mu_i| \tag{23}$$

If the test statistic exceeds a predefined threshold t_c the claim is rejected, otherwise the claimed identity is accepted, i.e.

$$\begin{array}{ll} d_c \leq t_c & accept\ \ claim \\ d_c > t_c & reject\ \ claim \end{array} \tag{24}$$

The threshold is chosen so as to achieve a specified operating point, that is a specified relationship between the false rejection of true claims and false acceptance of imposter claims. The operating point is determined from the *receiver operating characteristics* (ROC) curve which plots the relationship between these two error rates as a function of decision threshold. The ROC curve is computed on an independent set, known as *evaluation* set. Typically, the operating point is selected at the *equal error rate* (EER) where both the false rejection and false acceptance rates are the same.

Alternatively, one can pose the question how close the probe of the claimed identity is to the class of impostors, modelled by its mean $\mu_\Omega = -\frac{N_i}{N-N_i}\mu_i$. In this test we would expect the projected probe of a genuine claimant to be far from the projected imposter mean, giving the following test statistic:

$$\begin{array}{ll} d_i \leq t_i & reject\ \ claim \\ d_i > t_i & accept\ \ claim \end{array} \tag{25}$$

where

$$d_i = |\mathbf{a}_i \mathbf{z} - \frac{N_i}{N - N_i} \mathbf{a}_i \mu_i| \tag{26}$$

and t_i is the EER threshold obtained again from the ROC curve computed on the evaluation set. Note that for a sufficiently large training set the ratio of client to the total number of training images will be very small and the second term in (26) can be neglected. The imposter test statistics will then simply be the absolute value of the projection of the probe image into the client specific fisher face space.

In the Section 4 we present the results of experiments with both tests and show that in fact they are complementary. In other words they can be fused into a single decision scheme which exhibits superior performance.

3 Discussion

The proposed client specific LDA approach to personal identity verification has a number of interesting properties that are worth pointing out. First of all each client specific LDA representation space is only a one dimensional subspace of the original face space. This has important implications on the computational efficiency of the verification method. Commonly, LDA representation using common fisher faces would span a subspace of more than hundred dimensions. As the computational complexity of an LDA based method in the operation phase (after training) is linearly proportional to the LDA space dimensionality, the scheme developed will be more than hundred times faster than the conventional approach employing a common LDA basis. Moreover, as the test statistics are one dimensional, there is no need to compute a Euclidean distance in the fisher space. A simple comparison of a statistic (or in the worst case of its absolute value) with a threshold will yield a decision, thus achieving further computational gains. Finally, it is also worth noting that the projections of the client imposter means onto the client specific fisher face would be precomputed which would also significantly speed up the processing.

As far as the training phase is concerned, the proposed client specific LDA approach is also much simpler as each client specific fisher face, after the data has been projected into the PCA subspace, can be obtained without solving an eigenvalue analysis problem. This may not appear as such an advantage at a first glance as one still needs to find the inverse of the client specific within class scatter matrix Σ_i in (17) which is of similar complexity to eigenanalysis. However, for a sufficiently large number of training faces the mean scatter matrix M_i in (10) will tend to zero and its affect on Σ_i will be negligible. In fact in this case the within class scatter matrix will become common to all the clients,

i.e. $\Sigma_i = \Phi \quad \forall i$. Not only will the inverse of the within class scatter matrix become identical for all the clients. It will not have to be computed, as matrix Φ is nothing else but the matrix of eigenvalues of the mixture covariance matrix Σ. Hence it is a diagonal matrix and its inverse is obtained by inverting its diagonal elements. Thus finding the client specific fisher faces in this large training set case becomes trivial, avoiding computational intensive matrix operations. Most importantly, though, the determination of a client specific fisher face becomes independent of any other client specific fisher faces. This makes the enrollment of new clients particularly easy whereas in the conventional approach involving a common LDA space the introduction of a new client changes the scatter of means matrix and consequently all the client representations. This has serious implications on the practical utility and extensibility of the conventional approach in scenarios where the client population is continuously changing.

Apart from the merits of the proposed client specific LDA representation in terms of the speed of computation, the scheme has also very attractive properties from the point of view of the storage required. The actual advantages and drawbacks will depend the architecture of the personal identity verification system. In a fully centralised system a probe image of a claimant is transmitted to a remote central processing station which stores the details of each client. For the client specific LDA approach one needs to store $m \times (I + 1)$ data where I represents the size of the fisher face (image size). The conventional approach would require the storage of $m \times I$ PCAs plus $(m - 1) \times (m - 1)$ components of the LDA basis plus $m \times (m-1)$ template coefficients. Assuming the image size is an order of magnitude greater than the number of clients, both approaches will be comparable. However, when the number of clients approaches image size (number of image pixels), the proposed scheme will be slightly more efficient.

For a semicentralised scheme where some client specific data is stored on a smart card but the access claim processing is done in a local terminal storing the PCA bases, both approaches will be comparable. For the client specific approach one would need to store the client specific fisher face expressed in the PCA space (i.e. the $m-1$ dimensional

vector $\mathbf{v}_i$) where as for the conventional approach one would require to store the $m-1$ fisher face coefficients of the templates. In the fully localised processing where both the client information stored and the processing is performed in the smart card the client specific approach will be m time more efficient, both in terms of storage and processing speed.

One of the enhancements of the standard eigenface approach to personal identity verification and recognition is the idea of measuring the amount of probe image projection out of the face space. This information can be used as a feature to establish whether the probe image is consistent with the client face models. In the client specific approach one could still use this idea by implementing the mapping into the client specific fisher face in two stages. After the projection of the probe image into the PCA space one could check what proportion of the probe image lies outside this PCA space and reject the identity claim if this exceeds a certain threshold. The client specific fisher face verification would proceed in the next stage only if the initial test applied to the probe image is positive. However this approach would lose most of the computational advantages of the client specific LDA method. As it happens, with the two class formulation of the client specific personal identity verification problem the distribution of impostors has the mean close to the origin. As the image data is photometrically corrected before verification as described in Section 4, the variance of the grey level data and therefore the image magnitude is approximately constant. This means that if a significant proportion of the image vector lies outside the space, its projection into to the fisher space will also be foreshortened and the vector is likely to fall into the class of impostors. Thus in this two class case there is very little benefit in measuring the projection out of the PCA space explicitly.

In Section 2 we described a decision making strategy which involves two tests, client specific and imposter tests and their fusion to obtain the final decision. In theory, since the client and imposter distributions are projected into a one dimensional space, it should be possible to find a single threshold which separates the two classes at

given operating point error rates. This global threshold for EER has been found experimentally to derive an alternative decision strategy. Both schemes are experimentally compared in the following section.

As already alluded to, the enrollment of new clients in the proposed approach is considerably simpler than in the conventional approach, especially when one initially has sufficiently large database of clients to identify a representative PCA space. However, if the initial training set is relatively small it may be necessary to update the PCA axes and the client specific fisher faces as new clients are enrolled. This can be done by initially computing the PCAs recursively and eventually by applying the results of matrix perturbation analysis. Details of these approaches are beyond the scope of this paper.

4 Experiments

Experiments with the proposed method of personal identity verification using client specific fisher faces have been conducted on the XM2VTS face database [24] recorded at the University of Surrey. This publicly available database comprises video clips of 295 subjects recorded in four separate sessions over a period of five months. Within each session a number of shots were taken including two frontal view and head rotation sequences. In the frontal view sequence the subjects read aspecific text (providing synchronised speech and image data). In the rotation sequences both horisontal and vertical head movements were perfomed to provide information useful for 3D surface modelling of the head. The database of frontal face stills used in the experiments is only a small subset of the frontal face sequence frames. It was designated for algorithm benchmarking by the M2VTS consortium. The set contains 8 images (2 from each session).

The verification experiments were conducted according to the Lausanne protocol [24]. This protocol provides a standard framework for performance characterisation of personal identity verification algorithms so that the results of differnt methods are

directly comparable. The protocol specifies a partitioning of the database into three disjoint sets: a training set containing 200 clients, an evaluation set containing 200 clients and 25 impostors and a test set of 200 clients and 70 imposters. The imposter images in the evaluation and test sets are independent of each other and distinct from the client set. The training set is used to construct client specific fisher faces and models. The data in the vealuation set is used to determine verification thresholds and the test set is used to estimate the verification rate on independent data. Two configurations of the protocol, prescribing slightly different use of the training and evaluation data have been specified for experimentation.

The LDA approach is applicable only if all the images are correctly registered. Although automatic techniques for face image registration do exist, in our investigation of the advocated method we wanted to eliminate any contributory effects of misregistration on the performance of the method. For this reason we made use of manually determined eye coordinates for all the images in the database. The precision of this manual annotation is estimated to be within one pixel of the true eye positions. Using the manually detected eye coordinates each image was geometrically normalised to achieve registration.

Figure 1: Global mean, client mean and the client fisherface

Each image was then photometrically normalised by removing either just the image mean or by performing histogram equalisation. These normalisation procedures have been found most effective in similar studies involving conventional LDA representation, eigenface method and even support vector machines [16]. The global mean

of the registered and photometrically normalised images using histogram equalisation is shown in Figure 1. The mean of client ID no. 278 is shown in in the same figure, together with the client specific fisher face for client ID no. 278.

After training which first involved the determination of the PCA axes associated with the nonzero eigenvalues of the mixture covariance matrix and subsequently the computation of the client specific fisher faces, the evaluation set was used to compute the ROC curves. Three separate experiments were conducted to find the best strategy for photometric normalisation for each of the two formulations of the personal identity verification problem, i.e. testing for client acceptance and testing for imposter rejection. In the first experiment no normalisation was used whereas in experiments two and three the image mean subtraction and histogram equalisation were applied respectively. The corresponding ROC curves are shown in Figures 2 and 3. It is interesting to note that when testing for imposter rejection photometric normalisation significantly improves the performance, with histogram equalisation achieving the best results. Surprisingly, when testing for client claim acceptance, histogram equalisation was still the best photometric normalisation. However, the best results were obtained without any normalisation. This opens the possibility to use two diferent fisher faces to perform these two tests. However, in this report we have opted for a single client specific fisher face per client and chosen the normalisation by histogram equalisation which gave the best results overall.

Another view of the client/imposter separability can be gleaned from the histograms of the projections of the two classes, clients and imposters into the client specific fisherfaces. These are presented in Figures 4 and 5 respectively. Note that as expected all the imposter projections cluster at zero, i.e. the projected imposter mean. The client projections fall far from the origin. The negative projections are an artifact of the convention adopted for representing each fisher face. In principle, the client projections could be induced to be all positive as Section 2 suggests.

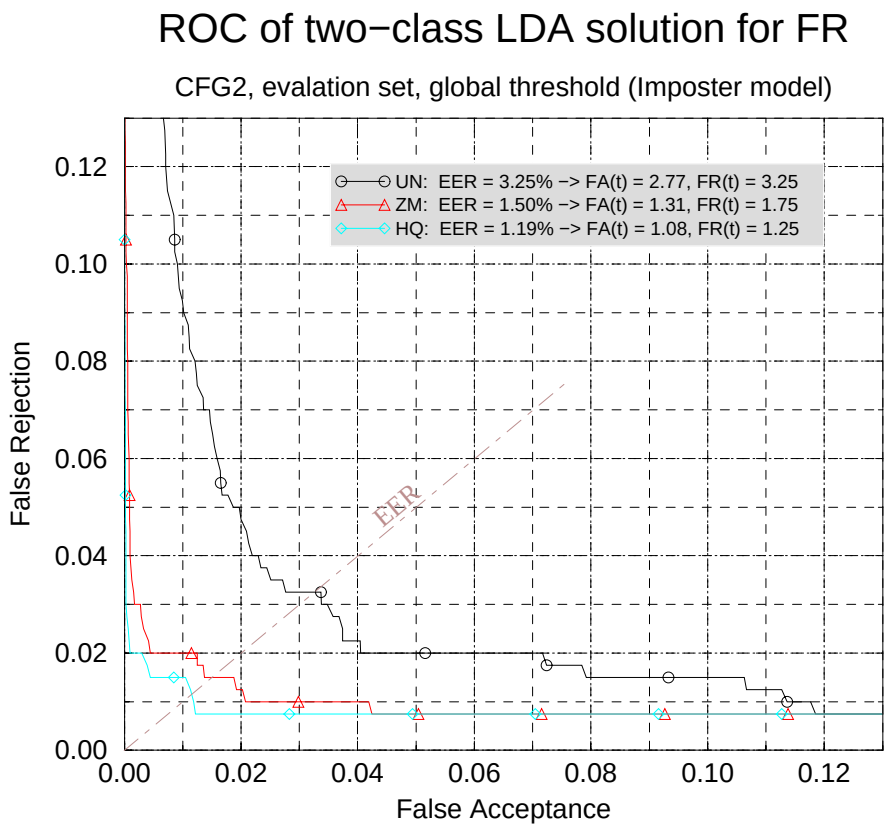

Figure 2: Performance against photometric normalisation: ROC on evaluation set using impostor-model

From the ROC curves the equal error rate thresholds were established for the two formulations of the identity authentication problem. This fully specified the verification method which was then tested on the independent test set. The results are shown in Table 1. The results show the false rejection, false acceptance and total error rates obtained on the evaluation set at the equal error threshold. The corresponding test set results are shown in the last three columns of the table.

For benchmarking purposes we have implemented a personal identity verification system based on the conventional fisher face approach with all clients sharing a common representation space. For 200 clients in the database this resulted in a 199 dimensional subspace of fisher faces. Each client was represented by its mean vector projected to this 199 dimensional subspace. An extensive experimental study of various scoring

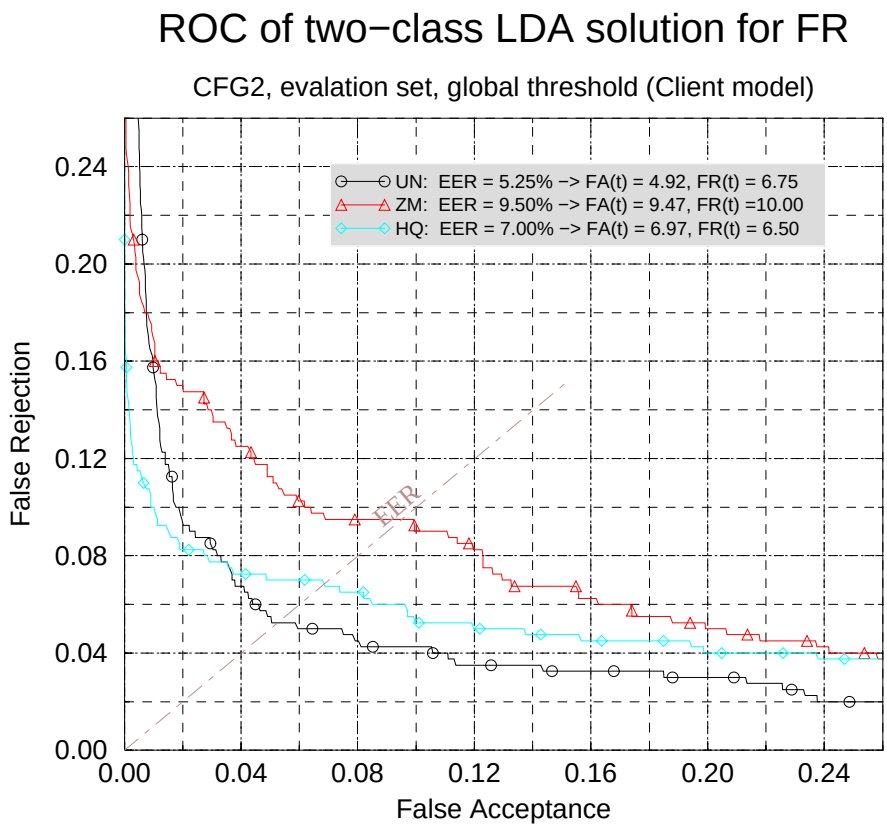

Figure 3: Performance against photometric normalisation: ROC on evaluation set using client-model

functions and threshold selection schemes which resulted in global and client specific thresholds were tested. The best results were obtained using normalised correlation in conjunction with client specific thresholds. The results are summarised in the rows of Table 1 marked Baseline.

Looking at Table 1 a number of observations can be made. In both configurations, better results were obtained using the imposter rejection test rather than the client acceptance test. However, a further improvement has been gained by fusing the results of these two tests. As we have only two experts, we could not opt for any sophisticated fusion strategy. As it turned out, a simple serial fusion scheme delivered a performance improvement as the decision errors were largely uncorrelated. More specifically, in the first stage of the decision system the claimed identity was tested for imposter rejec-

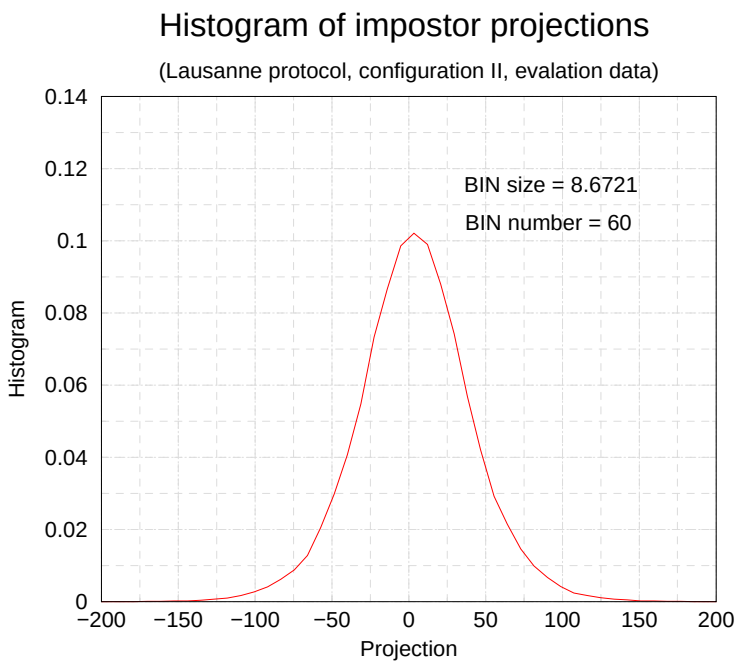

Figure 4: Histogram of impostors' projections

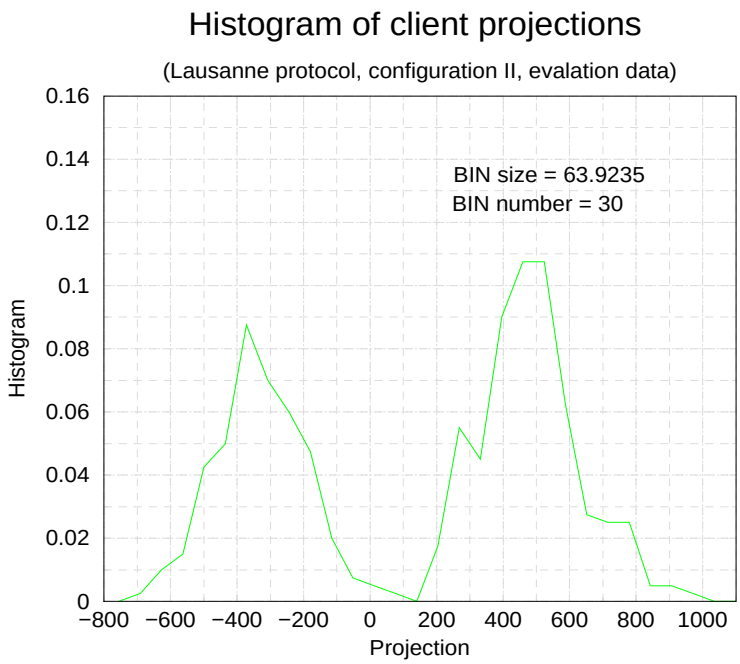

Figure 5: Histogram of clients' projections

CFG	Methods	Evaluation set			Test set		
		FR	FA	TE	FR	FA	TE
	Baseline	2.50	2.50	5.00	2.25	2.56	4.81
I	client-model	12.94	12.94	25.88	12.25	12.84	25.09
	impostor model	3.30	3.30	6.60	2.75	3.18	5.93
	fusion	2.67	0.70	3.37	2.25	1.14	3.39
	Baseline	1.25	1.25	2.50	0.25	1.23	1.48
II	client-model	7.00	7.00	14.00	6.50	6.98	13.48
	impostor model	1.19	1.19	2.38	1.25	1.08	2.34
	fusion	1.00	0.31	1.31	0.75	0.29	1.04

Table 1: Results comparison on XM2VTS database (image resolution 61×57)

tion. If the probe was rejected as an imposter, the claimed identity was accepted. If the imposter hypothesis was accepted then a second test, client acceptance test, had to be performed. If the client hypothesis was accepted the claimed identity would be accepted, else it would be rejected. The performance improvement gained by fusion are shown in the same table. For both configurations the results obtained are the best ever achieved with any of the methods and parameter setings we have experimented with so far. This includes not only conventional LDA method used as a baseline technique but also PCA, robust correlation and client specific Support Vector Machine. In the case of Configuration II the errors were more than halved.

5 Conclusion

In the paper we proposed one dimensional client specific fisher face representation. This novel LDA approach contrasts with the conventional LDA representation which involves multiple shared fisher faces. The method provides two measures for authentication: a distance to the client template, and a distance to the mean of impostors. These two decision scores are combined to achieve significant performance gains.

The method has been tested on the XM2VTS database according to the internationally agreed Lausanne protocol and shown to achieve superior performance. Interest-

ingly, the performance superiority is not the only advantage of the proposed method. Additional features of practical significance include the simplicity of training, as for large user databases the proposed technique requires only a matrix multiplication of the client mean vector. Moreover, the client enrollment is insulated from the enrollment of other clients. This opens the possibility to use other than the centralised architecture for the personal identity verification system and in fact smart card processing becomes a reality without any need to restrict the representation framework and therefore the representational capacity of the system. Finally the speed of probe testing is more than two orders of magnitude faster than that achieved by conventional PCA and LDA methods as the proposed techniques involves only a single fisher face per client. For these attractive properties the method is ideally suited for both representation and authentication in personal identity verification systems.

References

[1] P Belhumeur, J P Hespanha and D J Kriegman, "Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection," *IEEE trans. on PAMI*, Vol. 19, No. 7, pp. 711–720, July 1997.

[2] R Brunelli and T Poggio, "Face recognition: Features versus templates," *IEEE trans. onPAMI*, Vol. 15, pp. 1042-1052, 1993.

[3] R Chellappa, C L Wilson and S Sirohey, "Human and machine recognition of faces: A survey," *Proceedings of IEEE*, vol. 83, no. 5, pp. 705-740, 1995.

[4] T F Cootes, C J Taylor, D H Cooper and J.Graham, "Active Shape Models - Their Training and Application," *Computer Vision, Graphics and Image Understanding*, 61(1):38-59, 1995.

[5] I Craw, D Tock and A Bennett, Finding Face Features, Proc European Conference on Computer Vision, pp 92-96, 1992

[6] T Kanade, Computer Recognition of Human Faces, Birkhauser, Basel 1977

[7] P A Devijver and J Kittler, "Pattern Recognition: A Statistical Approach," Prentice-Hall, Englewood Cliffs, N.J., 1982.

[8] B Duc, E S Bigun, J Bigun, G Maitre and S Fischer, "Fusion of audio and video information for multi model person authent ication," *Pattern Recognition Letters*, Vol. 18, pp. 835-843, 1997.

[9] M Bichsel, Strategies of Robust Object Recognition for the Automatic Identification of Human Faces, PhD Thesis, ETH Zurich, Switzerland, 1991

[10] R Brunelli and T Pogio, Face Recognition through Geometrical Features, Proc. European Conference on Computer Vision, pp 792-800, 1992

[11] B Moghaddam and A Pentland, Probabilistic Visual Learning for Object Detection, in Proc. International Conference on Computer Vision, pp 786-793, 1995

[12] G J Edwards, T F Cootes and C J Taylor, "Face Recognition using Active Appearance Models," In *Proc. of ECCV'98*, pp. 581-595, 1998.

[13] K Etemad and R Chellappa, "Face Recognition using Discriminant Eigenvectors," In *Proc. of ICASSP'96*, pp. 2148-2151, 1996.

[14] K Etemad and R Chellappa, "Discriminant Analysis for Recognition of Human Face Images," *Journal of Optical Society of American*, pp. 1724-1733, August, 1997.

[15] K Jonsson, G Matas, and J Kittler, "Learning Salient Features for Real-Time Face Verification," In *Proc. of AVBPA'99*, pp. 60-65, 1999.

[16] K Jonsson, J Kittler, Y P Li and J Matas "Support Vector Machine for Face Authentication," In *Proceeding of MMVC'99*, pp. 543-553. 1999.

[17] M Kirby, and L Sirovich, "Application of the Karhunen-Loeve Procedure for the Characterisation of Human Faces," *IEEE tran. on PAMI*, Vol. 12, No. 1, Jan. 1990.

[18] J Kittler, M Hatef, R P W Duin and J Matas, "On Combining Classifiers," *IEEE trans. on PAMI*, Vol. 20, No. 3, pp. 226-239, March 1998.

[19] C Kotropoulos, A Tefas, I Pitas and C Fernandez, "Performance Assessment of Morphological Dynamic Link Architecture under Optimal and Real Operating Conditions," in *Proc. of 1998 IEEE Int. Conf. on Image Processing*, Chicago, 1998.

[20] M Lades, J C Vorbruggen, J Buhmann, J Langem C von Malsburg, R P Wurtz and W Kohen, "Distortion invariant object recognition in the dynamic link architecture," *IEEE Trans. on Computers*, Vol. 42, no. 3, pp. 300-311, March 1993.

[21] Y P Li, J Kittler and J Matas, "Effective implementation of Linear Discriminant Analysis for face recognition and verification," In the Proceeding of CAIP'99, pp. 234–242, 1999.

[22] C Liu and H Wechsler, "Face Recognition Using Evolutionary Pursuit", in *Proc. of ECCV'98*, Vol. II, pp. 596-612, June 1998.

[23] J Matas, K Jonsson and J Kittler, "Fast Face Localisation and Verification," In Proceeding of BMVC'97, pp. 152–161, 1997.

[24] K Messer, J Matas, J Kittler, J Luettin and G Maitre, "XM2VTSDB: The Extended M2VTS Database," In *Proc. of AVBPA'99*, pp.72-77, 1999

[25] B Moghaddam and A Pentland, "Probabilistic visual learning for object representation," In S.K. Nayer and T. Poggio,editors, *Early Visual learning*, pp. 99-130, 1997.

[26] B Moghaddam, W Wahid and A Pentland, "Beyond Eigenfaces: Probabilistic Matching for Face Recognition," In *proc. of Face and Gesture'98*, pp. 30-35, 1998.

[27] H Moon and P J Philips, "Analysis of PCA-based Face Recognition Algorithms," In proceeding of FG'98, pp. 205-210.

[28] A Pentland, B Moghaddam and T Starner, "View-based and Modular Eigenspaces for Face Recognition," in *Proc. of CVPR'94*, pp. 84-91.

[29] "Support Vector Machines Applied to Face Recognition," In M. S. Kearns et al, editors, *Advances in Neural Information Processing Systems 11* , 1998.

[30] S Romdhani, "Face Recognition using Principal Component Analysis," MSc Thesis, Univ. of Glasgow. *http://www.elec.gla.ac.uk/~romdhani/pca_doc/pca_doc_toc.htm*.

[31] L Sirovich and M Kirby, "Low-dimensitional procedure for the characterization on human faces," *J.Opt.Soc. Am. A*, Vol. 4, no. 3, pp 519-524, 1987.

[32] D L Swets and J Weng, "Discriminant Analysis and Eigenspace Partition Tree for Face and Object Recognition from Views," In proceeding of FG'96, pp. 192-197, Killington, Vermont, Oct. 14-16, 1996.

[33] D L Swets and J Weng, "Hierarchical Discriminant Analysis for Image Retrieval," *IEEE Tran. on PAMI*, vol. 21, no. 5, pp. 386-401, 1999.

[34] M Turk and A Pentland, "Eigenface for Recognition," *Journal of Cognitive Neuroscience*, Vol. 3, no. 1, pp. 70-86, 1991.

[35] K Fukunaga, Intorduction to Statistical Pattern Recognition, Academic Press, New York 1990

[36] W Zhao et al, "Discriminant Analysis of Principal Components for Face Recognition," In Harry Wechsler et al,editors,*Face Recognition From Theory to Applications*, NATO ASI Series, pp. 73-85, Springer, 1998.