Available online at www.sciencedirect.com

ScienceDirect

Image and Vision Computing 25 (2007) 553–563

www.elsevier.com/locate/imavis

Gabor wavelets and General Discriminant Analysis for face identification and verification

LinLin Shen [a,*], Li Bai [a], Michael Fairhurst [b]

[a] School of Computer Science & IT, University of Nottingham, UK
[b] Electronics Department, University of Kent, UK

Received 14 June 2004; received in revised form 3 April 2006; accepted 16 May 2006

Abstract

A novel and uniform framework for both face identification and verification is presented in this paper. The framework is based on a combination of Gabor wavelets and General Discriminant Analysis, and can be considered appearance based in that features are extracted from the whole face image. The feature vectors are then subjected to subspace projection. The design of Gabor filters for facial feature extraction is also discussed, which is seldom reported in the literature. The method has been tested extensively for both identification and verification applications. The FERET and BANCA face databases were used to generate the results. Experiments show that Gabor wavelets can significantly improve system performance whilst General Discriminant Analysis outperforms other subspace projection methods such as Principal Component Analysis, Linear Discriminant Analysis, and Kernel Principal Component Analysis. Our method has achieved 97.5% recognition rate on the FERET database, and 5.96% verification error rate on the BANCA database. This is a significantly better performance than that attainable with other popular approaches reported in the literature. In particular, our verification system performed better than most of the systems in the 2004 International Face Verification Competition, using the BANCA face database and specially designed test protocols.

Keywords: Face identification; Face verification; Gabor wavelets; General Discriminant Analysis

1. Introduction

Automatic face recognition has rapidly developed over the years and is now a highly active field of research, with important applications in security surveillance, access control, human–machine interaction, and a host of other domains.

Early face recognition work was represented by Von der Malsburg's Dynamic Link Architecture [1], an elastic matching process in which a test face, represented as a graph with Gabor wavelet responses as the nodes, is associated with one of a number of stored face graphs. The matching process maximises the similarities between the test and the corresponding model face graph. A variant of this is the Face Bunch Graph [2], which was proposed to cope with the specific variability of face images. Gabor wavelets are applied at manually selected fiducial points (eyes, mouth, nose, etc.) of several images of a face and the results, referred to as "jets", are packed in a graph called the face bunch graph. Disadvantages of this approach include rigid alignment of facial features and limited ability in handling pose variations. Instead of graphs, Hidden Markov Models (HMMs) represent a face image as a sequence of 'states'. The goal of training a HMM is to optimise its parameters to 'best' describe the observation vectors representing a class, and recognition is carried out by matching a test image against each of the trained HHMs. An interesting variant of HMM is the Embedded HMM [3], in which super-states of an embedded HMM represent primary facial regions (forehead, eyes, nose,

* Corresponding author. Tel.: +44 115 8466567; fax: +44 115 951 4254.
E-mail addresses: lls@cs.nott.ac.uk (L. Shen), bai@cs.nott.ac.uk (L. Bai), m.c.fairhurst@kent.ac.uk (M. Fairhurst).

0262-8856/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.imavis.2006.05.002

mouth, and chin) whilst embedded states within each super-state describe in more detail each of the facial regions. This approach is computationally expensive, as it requires the computation of a probabilistic distance to each stored HHM when classifying a new image.

Linear transform based methods such as Eigenface [4], Fisherface [5], and Independent Component Analysis (ICA) [6] have had a significant influence within the face recognition community for a considerable time. Eigenfaces are a set of eigenvectors arising from applying Principal Component Analysis (PCA) to a collection of images, and any face image could then be described by its projection coefficients onto the eigenfaces thus generated. This significantly reduces the dimension of the relevant face vectors, giving greater tractability in practical application. Fisherfaces are based on Linear Discriminant Analysis (LDA), the objective of which is to maximize class separability, defined as the ratio of the between-class scatter matrix to the within-class scatter matrix. Whilst in PCA the emphasis is on de-correlation of variables, ICA aims at variable independence, a much stronger condition. However, Baek [7] presented a comparison of PCA and ICA and concluded that, when a proper distance metric was used, PCA outperformed ICA significantly on the FERET face database [8] of more than 1000 images. Kernel based methods [9], exemplified by Kernel Principal Component Analysis (KPCA) [10,11], Kernel Fisher Discriminant Analysis (KFDA) [12,13] and General Discriminant Analysis (GDA) [14] have significantly outperformed PCA, LDA, ICA and neural networks in similar recognition tasks. The support vector machine (SVM) approach [15] is another example of a Kernel method, which seeks a unique hyperplane yielding the maximum margin of separation between two classes through constrained optimisation. Phillips [16] trained a single SVM to distinguish between within-class and between-class images. Jonsson [17] trained a SVM for each class. Both used linear transform based methods for feature extraction.

Although significant progress has been made in appearance-based face recognition, to our knowledge, the application of KPCA or GDA in face verification has seldom been reported in the literature. Despite the wide application of Gabor filters for feature extraction [1,2,18], the design of Gabor filters is rarely discussed. In this paper, we describe an approach that combines Gabor feature extraction and Kernel subspace projection techniques to produce a robust and uniform system framework for both face identification and verification. This involves extracting discriminant features from images using a sequence of Gabor wavelets at different scale and orientations and projecting feature vectors to a subspace before identification, or verification, can take place. The design of Gabor filters is also discussed based on experiments. Our method is hereafter referred to in this paper as the Gabor + GDA.

Extensive experiments have been conducted to evaluate the performance of Gabor + GDA against existing methods in the literature. Gabor features have also shown robustness against variations of head pose and camera orientation The generalization ability of the novel Gabor + GDA method has also been observed. Since the FERET [8] and BANCA database [19,20], specially designed to test face identification and verification algorithms, are used for the evaluation of our algorithms, our results are directly comparable with other methods, and comparisons with a number of popular approaches will be made to illustrate the advantages of the proposed new method.

The contribution of this paper is therefore threefold. First, we discuss how to design Gabor filters empirically for facial feature extraction and demonstrate that the proposed novel Gabor + GDA framework is robust and uniform for both identification and verification. Second, we show that GDA outperforms other subspace projection techniques such as PCA, LDA, and KPCA, and that different distance measures can have significant effects on subspace based methods. Finally, we show evidence to support the findings reported by some other researchers that PCA outperforms LDA when the training set is non-representative.

The paper is organized as follows. In Section 2, Gabor wavelets are defined and the methodology to extract discriminative Gabor features from face images are outlined. Section 3 introduces Generalized Discriminant Analysis, while the strategy to combine the Gabor and GDA concepts is given in Section 4. Experimental results for identification and verification using the FERET and BANCA database are given in Section 5, and some important conclusions are drawn in Section 6.

2. Gabor feature extraction

2.1. Gabor wavelets

The characteristics of the Gabor wavelets (filters), especially for frequency and orientation representations, are similar to those of the human visual system, and they have been found to be particularly appropriate for texture representation and discrimination. The Gabor filters-based features, directly extracted from gray-level images, have been successfully and widely applied to texture segmentation [21,22], handwritten numerals recognition [23] and fingerprint recognition [24]. In the spatial domain, a 2D Gabor filter is a Gaussian kernel function modulated by a sinusoidal plane wave:

$$\begin{aligned} \varphi_{\Pi(f,\theta,\gamma,\eta)}(x,y) &= \frac{f^2}{\pi\gamma\eta}\exp(-(\alpha^2x'^2+\beta^2y'^2))\exp(j2\pi fx') \\ x' &= x\cos\theta + y\sin\theta \\ y' &= -x\sin\theta + y\cos\theta \end{aligned} \tag{1}$$

where f is the central frequency of the sinusoidal plane wave, θ is the anti-clockwise rotation of the Gaussian and the plane wave, α is the sharpness of the Gaussian along the major axis parallel to the wave, and β is the sharpness of the Gaussian minor axis perpendicular to the wave. $\gamma = \frac{f}{\alpha}$ and $\eta = \frac{f}{\beta}$ are defined to keep the ratio between frequency and sharpness constant. The Gabor filters are self-similar – all filters can be generated from one mother wavelet by dilation and rotation. Each filter is in the shape of plane waves with frequency f, restricted by a Gaussian envelope function with relative width α and β. To extract useful features from an image, e.g. face, normally a set of Gabor filters with different frequencies and orientations are required:

$$\varphi_{u,v} = \varphi_{\Pi(f_u,\theta_v,\gamma,\eta)}, \quad f_u = f_{\max}/\sqrt{2}^u, \quad \theta_v = \frac{v}{8}\pi,$$
$$u = 0,\ldots,U-1, v = 0,\ldots,V-1 \qquad (2)$$

2.2. Design of Gabor filters for facial feature extraction

As shown in Eqs. (1) and (2), following parameters need to be determined to design Gabor filters for feature extraction: the highest peak frequency $f_{\max}$, the ratio between centre frequency and the sharpness of Gaussian major axis: γ and minor axis: η, the number of scales U and orientations V. According to the Nyquist sampling theory, a signal containing frequencies higher than half of the sampling frequency cannot be reconstructed completely. Therefore, the upper limit frequency for a 2D image is 0.5 cycles/pixel, while the low limit is 0. However, useful frequency band of face images is much narrower. We observe from sample face images in the frequency domain that useful information is mainly contained in the low frequency band. Therefore, we choose $f_{\max} = 0.25$ for face recognition. Parameters γ and η determine the ratio between the centre frequency and the size of Gaussian envelope. Once the ratio is fixed, the size of the Gaussian envelope monotonically decreases with the value of centre frequency. The higher is the centre frequency of the Gabor sinusoidal carrier, the smaller area the Gaussian envelop will cover in spacial domain. This is reasonable since the high frequency signal changes faster. We assume $\alpha = \beta$ and $\gamma = \eta = \sqrt{2}$, as used by most researchers [2,18]. Since there is no theoretical basis available for selecting scales and orientations, we performed some experimental evaluations on different combinations of the two parameters. The results show that 5 scales and 8 orientations shall be used for face recognition purpose. Details of the experiments are shown in Section 5.1.

2.3. Gabor representation of face images

The Gabor representation of a face image $I(\vec{x})$ can be obtained by convolving the image with the family of Gabor filters as defined by (1):

$$G_{u,v}(\vec{x}) = (I * \varphi_{u,v})(\vec{x}) \qquad (3)$$

where $G_{u,v}(\vec{x})$ denotes the convolution result corresponding to the Gabor filter at orientation u and scale v. Fig. 1. shows the convolution result of a face image with two Gabor filters. The face image is convolved with two Gabor filters at different orientations and scales, both magnitude and real part of the convolution results are shown. As a result, image $I(\vec{x})$ can be represented by a set of Gabor wavelet coefficients $\{G_{u,v}(\vec{x}), u = 0,\ldots,4; v = 0,\ldots,7\}$. The magnitude of each $G_{u,v}(\vec{x})$ is then downsampled by a factor r, normalized to zero mean and unit variance, and turned to a vector $x^r_{u,v}$ by concatenating the rows [18]. A discriminative feature vector x^r can be derived to represent the image I by concatenating those vectors $x^r_{u,v}$:

$$x^r = ((x^r_{0,0})^t(x^r_{0,1})^t \cdots (x^r_{4,7})^t)^t \qquad (4)$$

The derived feature vector x^r thus encompasses all the elements of the Gabor wavelet representation set $\{G_{u,v}(\vec{x}), u = 0,\ldots,4; v = 0,\ldots,7\}$. However, the dimension of the vector is quite high, e.g., for a 128×128 image, the vector dimension is $l = 10240$ when the downsampling factor $r = 64$. In the following sections, we will introduce subspace projection techniques to reduce feature dimension.

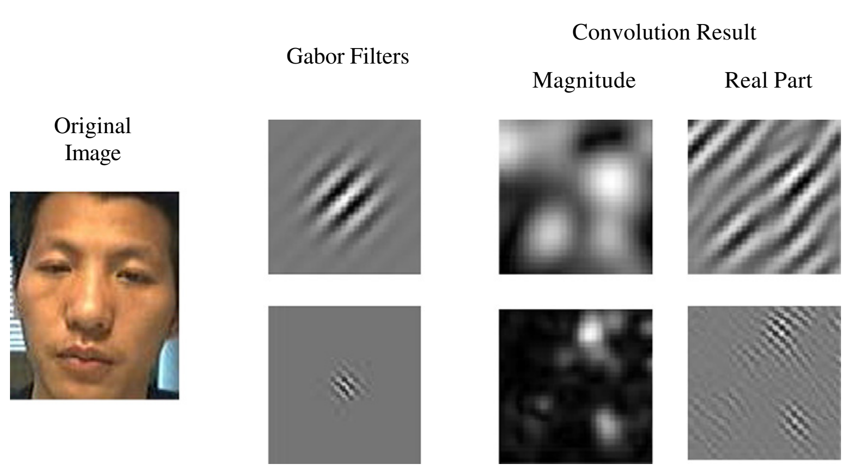

Fig. 1. Convolution results of a face image with two Gabor filters.

3. Generalized Discriminant Analysis

Classical Discriminant Analysis techniques, such as Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA), have been successfully applied to face recognition since Turk's pioneering work [4]. Well known as the Eigenface method, PCA identifies a subspace spanned by the training images, which could decorrelate the variance of pixel values. Unlike PCA, LDA aims to find a projection matrix W which is optimized to separate different classes, i.e., maximizes the quotient of the determinant of S_b and S_w,

$$W = \arg\max \frac{|W^T S_b W|}{|W^T S_w W|} \tag{5}$$

where S_b, S_w are the between-class scatter and within-class scatter, respectively. It was shown in [25] that the projection matrix W can be computed from the eigenvectors of $S_w^{-1}S_b$. However, due to the high dimensionality of the feature vector, especially in face recognition, S_w is usually singular, i.e., the inversion of S_w does not exist. Several techniques, such as the PCA + LDA [26], Regularized LDA (RLDA) [27], Enhanced LDA (ELDA) [28] and Direct LDA (DLDA) [29] have been proposed in literature to solve the Small Sample Size (SSS) problem.

Similar to LDA, the purpose of GDA [14] is to maximize the quotient between the inter-classes inertia and the intra-classes inertia in a mapped feature space. Considering a C-class problem and letting N_c be the number of samples in class c, a set of training patterns from the C classes can be defined as $\{x_{ck},\ c = 1, 2, \ldots, C;\ k = 1, 2, \ldots, N_c\}$, $N = \sum_{c=1}^{C} N_c$. Given a nonlinear mapping ϕ: $R^N \to F$, the set of training samples in the mapped feature space can be represented as $\{\phi(x_{ck}),\ c = 1,\ 2, \ldots, \mathrm{C};\ k = 1,\ 2, \ldots,\ N_c\}$. The S_b and S_w of the training set can be computed as:

$$S_w = \frac{1}{C}\sum_{c=1}^{C}\frac{1}{N_c}\sum_{k=1}^{N_c}\phi(x_{ck})\phi(x_{ck})^T \tag{6}$$

$$S_b = \frac{1}{C}\sum_{c=1}^{C}(\mu_c - \mu)(\mu_c - \mu)^T \tag{7}$$

GDA finds the eigenvalues $\lambda \geqslant 0$ and eigenvectors $v \in F\backslash\{0\}$ satisfying

$$\lambda S_w v = S_b v \tag{8}$$

where all solutions $\mathbf{v}$ lie in the span of $\phi(x_{11}), \ldots, \phi(x_{ck}), \ldots$ and there exist coefficients α_{ck} such that

$$v = \sum_{c=1}^{C}\sum_{k=1}^{N_c}\alpha_{ck}\phi(x_{ck}) \tag{9}$$

Using kernel techniques, the dot product of a sample i from class p and the other sample j from class q in the feature space, denoted as $(k_{ij})_{pq}$, can be calculated by a kernel function, e.g., radial basis kernel as below:

$$(k_{ij})_{pq} = \phi(x_{pi})\cdot\phi(x_{qj}) = k(x_{pi}, x_{qj}) = e^{-|x_{pi}-x_{qj}|^2/r} \tag{10}$$

Let $\mathbf{K}$ be a $M \times M$ matrix defined on the class elements by $((\mathbf{K}_{pq})_{\substack{p=1,\ldots,C\\ q=1,\ldots,C}})$, where $\mathbf{K}_{pq}$ is a matrix composed of dot products between vectors from class p and q in feature space:

$$\mathbf{K}_{pq} = (k_{ij})_{\substack{i=1,\ldots,N_p\\ j=1,\ldots,N_q}} \tag{11}$$

We also define a $M \times M$ block diagonal matrix:

$$\mathbf{U} = (\mathbf{U}_c)_{c=1,\ldots,C} \tag{12}$$

where $\mathbf{U}_c$ is $N_c \times N_c$ a matrix with terms all equal to $\frac{1}{N_c}$.

By substituting (6), (7) and (9) into (8) and taking inner-product with vector $\phi(x_{ij})$ on both sides, the solution of (8) can be achieved by solving:

$$\lambda\mathbf{KK}\ \mathbf{a} = \mathbf{KUK}\ \mathbf{a} \tag{13}$$

where $\mathbf{a}$ denotes a column vector with entries α_{ck}, $c = 1, \ldots, C$, $k = 1, \ldots, N_c$. The solution of a in Eq. (13) is equivalent to find the eigenvectors of the matrix $(\mathbf{KK})^{-1}$ $\mathbf{KUK}$. However, similar to the SSS problem, the matrix $\mathbf{K}$ might not be reversible. We find the eigenvector $\mathbf{a}$ by first diagonalising matrix K (see [14] for more details). Once the first L significant eigenvectors are found, a projection matrix can be constructed as:

$$\mathbf{W} = [\mathbf{a}_1\mathbf{a}_2\ldots\mathbf{a}_L] \tag{14}$$

The projection of x in the L-dimensional GDA space is given by:

$$\mathbf{y} = \mathbf{k_x}\mathbf{W} \tag{15}$$

where

$$\mathbf{k_x} = [k(x, x_{11})\ldots k(x, x_{ck})\ldots k(x, x_{CN_c})] \tag{16}$$

4. Combining Gabor feature and GDA for identification and verification

In our approach we propose to transform the Gabor feature space into the GDA space to reduce Gabor feature dimension for face recognition. Compared with the raw pixel features used by classical subspace methods such as Eigenface and Fisherface, Gabor features contain more discriminant information and are thus more robust against variations in illumination, pose and expressions. Figs. 2 and 3 show the system structure for face identification and verification respectively. The identification task is inherently more difficult than verification, since the input face image must be compared with, and matched to, each face in the enrolment database. The test face is then identified as belonging to the face class (i.e. the individual) which shows the highest similarity.

Both identification and verification require a similarity measure reflecting the differences between two facial features in the GDA space. Given a Gabor feature vector extracted from a face image x and a subspace projection matrix $\mathbf{W}$ derived from the GDA subspace analysis, a discriminant feature y with low dimension can be derived by $\mathbf{y} = \mathbf{k_x}\mathbf{W}$. As described below, three different distance

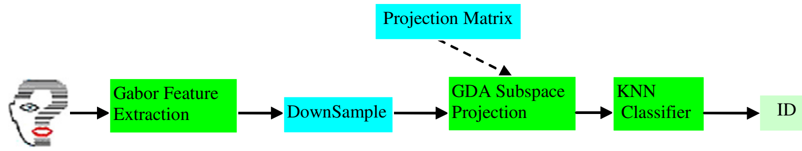

Fig. 2. Gabor + GDA for identification.

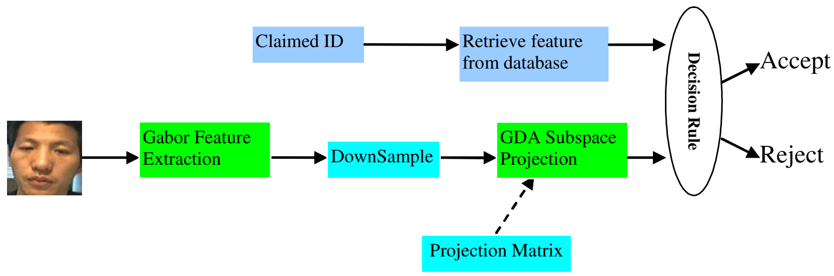

Fig. 3. Gabor + GDA for verification.

measures d_E, d_c, d_M are used in our work to calculate the distance between two sample projections y_1 and y_2. These are:

Euclidean Distance (Eu):

$$d_E(y_1, y_2) = \sqrt{(y_1 - y_2)^T (y_1 - y_2)} \quad (17)$$

Mahalanobis Distance (Ma):

$$d_M(y_1, y_2) = (y_1 - y_2)^T \sum^{-1} (y_1 - y_2) \quad (18)$$

Normalized Correlation (Nc):

$$d_C(y_1, y_2) = \frac{y_1^T y_2}{\|y_1\| \|y_2\|} \quad (19)$$

where $\sum$ is the covariance matrix, and $\|\cdot\|$ denotes the norm operator. While the simple nearest neighbour classifier is used in our work for identification, the decision rule in verification is defined as follows: the claimed identity is accepted if the mean distance of the test face with the training images of the identity is below a global threshold, otherwise it is rejected.

5. Experimental results

5.1. Results for identification

5.1.1. The FERET database

The standard test bed adopted in similar studies, the FERET database [8], was used to test our algorithm for identification. The results are reported in terms of the standard recognition rate. Specifically, this is the ratio between the number of correctly identified face images and the number of test images. Six hundred frontal face images from 200 subjects are selected, where all the subjects are in an upright, frontal position, with tolerance for some tilting and rotation of up to 10 degrees. The 600 face images were acquired under varying illumination conditions and facial expressions. Each subject has three images of size 256×384 with 256 gray levels. The following procedures were applied to normalize the face images prior to the experiments:

- The centers of the eyes of each image are manually marked, each image is rotated and scaled to align the centers of the eyes.
- Each face image is cropped to the size of 128×128 to extract the facial region, and normalized to zero mean and unit variance.

To test the algorithms, two images of each subject are randomly chosen for training, while the remaining one is used for testing. Fig. 4 shows sample images from the database. The first two rows are the example training images while the third row shows the example test images. It can be seen from this figure that the test images all display variations in illumination and facial expression.

5.1.2. Tuning Gabor filter parameters

In the first experiment, we aim to find the best Gabor filter parameters for face recognition purposes. The simple PCA subspace projection and the nearest neighbour classifier are used to test the parameters. The number of scales of Gabor filters defines the range of frequency information the filters can extract, while the number of orientations specifies the resolution of the extracted directional features. The larger is the number of scales, the more information from low frequency bands will be extracted. We first vary the number of orientations and keep the number of scales fixed Fig. 5 shows the performance of the face recognition system when number of orientations varies between 5 and 3. It can be observed from the figure that the performance increase with the number of orientations. The higher is the orientation resolution, the better the performance. The results suggest that important features in faces are in

Fig. 4. Example training images (top 2 rows) and test images (bottom row) of the FERET database.

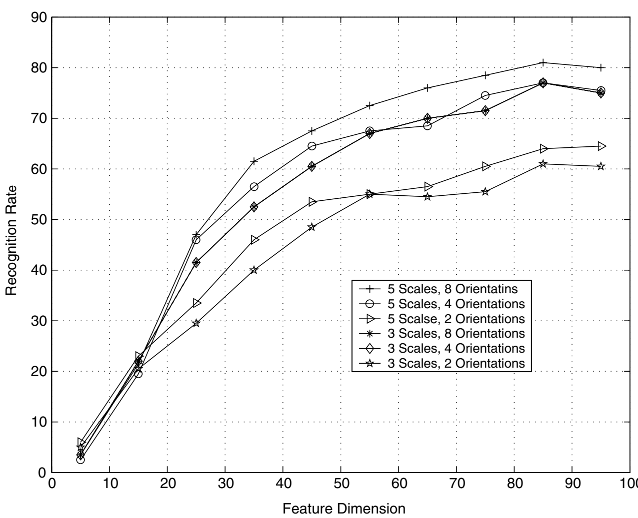

Fig. 5. Performance for different number of orientations.

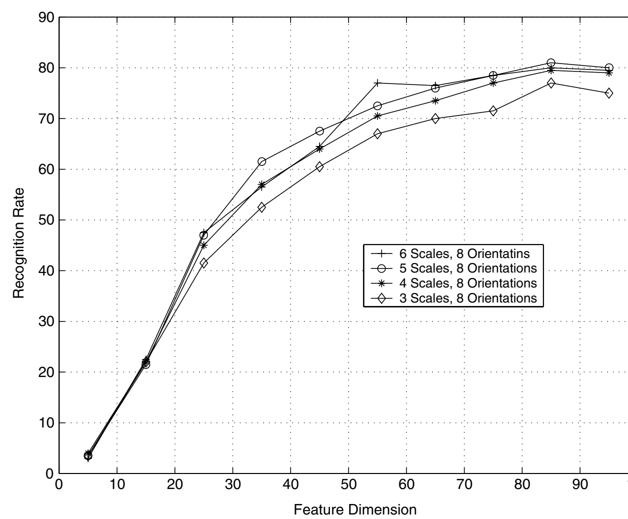

Fig. 6. Performance for different number of scales.

different orientations, e.g., eyes, mouths, etc. Eight orientations seem to be a good choice in our experiments.

The system initially performs better as the number of scales of Gabor filters increases. However, when the number of scales reaches five, the performance becomes stable, see Fig. 6. Since most of the useful information in face images is contained within a limited frequency band, the inclusion of more scales will result in redundant information and thus reduce system performance. Five scales appear to be have achieved good performance in our experiments. Therefore, we choose to use Gabor filters of five scales and eight orientations in the following experiments.

5.1.3. Comparison of different distance measures

In the next experiment, we tested with the three similarity measures Eu, Ma and Nc for the proposed Gabor + GDA method. As shown in Fig. 7, normalized correlation achieved the best performance for GDA among the three distance measures, while the difference between performance using Euclidean distance and Mahalanobis

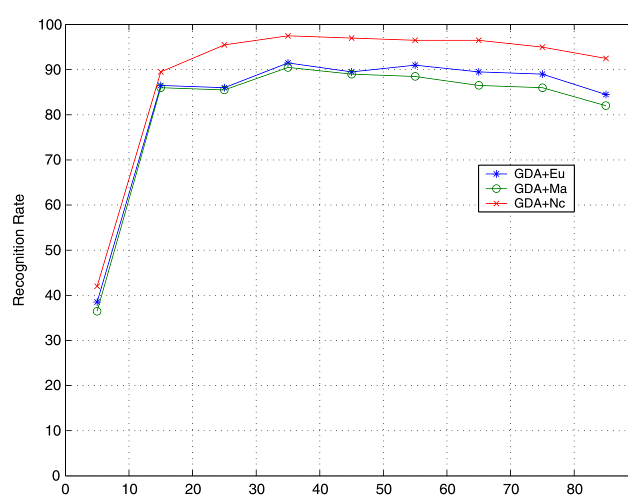

Fig. 7. Performance of Gabor + GDA using different distance measures.

distance is not large. However, the results are different for different subspace projection methods. For example, Mahalanobis is the best distance measure for KPCA, which achieves significantly higher recognition rates than the other two measures. Similar results are also observed for the linear subspace projection methods, PCA and LDA. It seems that for expressive features derived in PCA and KPCA space, the Mahalanobis distance measure is more suitable than the others; while for discriminating features extracted by LDA and GDA, the correlation distance measure seems to be the best choice.

5.1.4. Comparison with other subspace methods

The comparative results of PCA, LDA, KPCA and GDA on the Gabor feature vector with the respective optimized distance measures are shown in Fig. 8.

It can be seen that nonlinear subspace methods are performing fundamentally better than their corresponding linear approaches. In other words, KPCA performs better than PCA and GDA performs better than LDA. In fact, GDA performs the best among these four algorithms while, following GDA, LDA performs better than KPCA and PCA. A recognition rate as high as 97.5% is achieved for the novel Gabor + GDA approach when the dimension is set at 35.

5.1.5. Gabor features versus raw pixel values

To emphasise the discriminating power of the extracted Gabor feature vector, the comparative performance of PCA, Gabor + PCA, GDA and Gabor + GDA are also shown in Fig. 9. When the Gabor feature vector is not used, the pixel values of face images are simply concatenated to a feature vector. For example, the length of a raw pixel feature vector will be $128 \times 128 = 16{,}384$ for an image with size 128×128. It is apparent that the adoption of the Gabor feature vector improves the performance of PCA and GDA by a large margin. The Gabor + PCA method achieves 20% higher accuracy than PCA, while 6% improvement is observed for GDA when Gabor wavelets are applied. An improvement for Gabor + LDA and Gabor + KPCA has also been observed in the experiments.

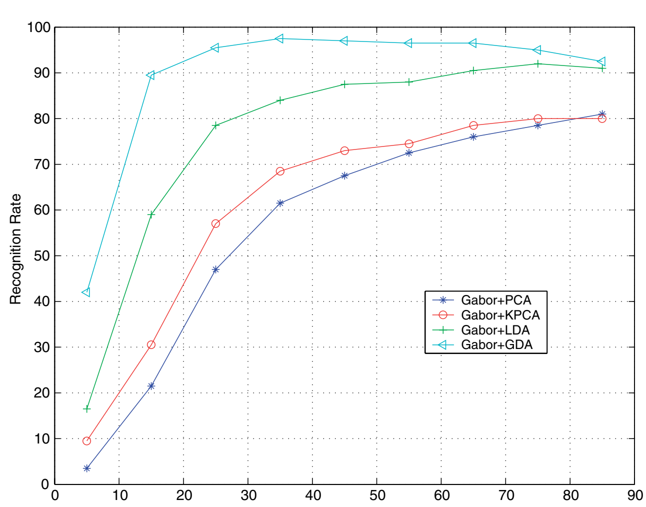

Fig. 8. Experimental results of PCA, LDA, KPCA and GDA using Gabor features.

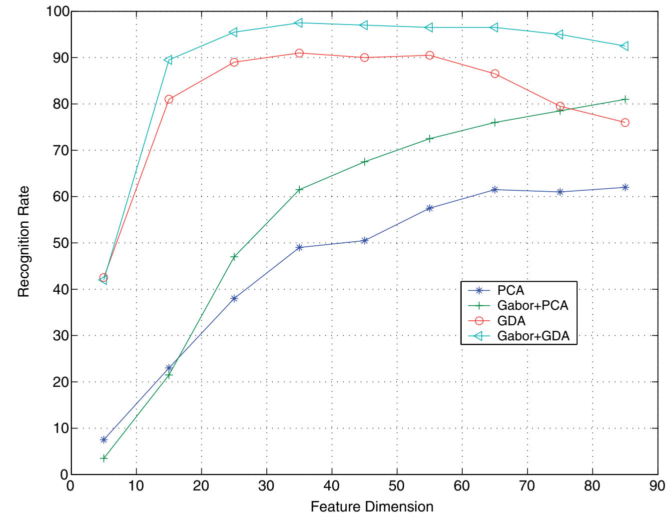

Fig. 9. Performance improvement of PCA and GDA using Gabor features.

5.1.6. Comparison with other methods

For further comparison of Gabor feature and GDA based methods with other approaches, the results on the same database for the Radial Basis Function (RBF) neural network [30] and HMM [3,31] based methods are shown in Table 1. Raw pixel features are used for RBF based methods, in that the normalized pixel values of the image are input directly to the network for personal identity determination. The two layers of the RBF network and HMMs are trained using the same training set, with parameters optimized for best performance. To make RBF the same structure as in the case of GDA, which takes the inner product of the input data with all of the training samples, the network is designed with 400 nodes for the input layer and 200 nodes for the output layer. While the DCT-HMM uses DCT (Discrete Cosine Transform) coefficients for observation vector extraction, DWT-HMM adopts DWT (Discrete Wavelet Transform) for more robust feature extraction. As shown in Table 1, Gabor + GDA performs significantly better than either of the other two methods used for comparison.

Table 1
Comparative results of Gabor + GDA with other methods on part of the FERET database

Method	Recognition rate (%)
RBF Network	75
DCT-HMM	32.5
DWT-HMM	44.5
Gabor + GDA	97.5

5.2. Results for verification

5.2.1. The BANCA database

The BANCA database [19] is used to test our algorithms in the corresponding verification application, which has been specially designed to test face verification systems. The BANCA database consists of images from 52 subjects captured in 12 sessions. Ten face images are captured for each person in each session. The 12 sessions are composed of 3 different scenarios: (1) Controlled scenario (c) for sessions 1–4, (2) Degraded scenario (d) for sessions 5–8, (3) Adverse scenario (a) for sessions 9–12. A web cam was used in the Degraded scenario and a high quality camera was used in the Controlled and Adverse scenarios. Images are captured with normal pose in the Controlled and Adverse scenarios, whilst the head down pose is required in the Adverse scenario.

Fig. 10 shows the sample images captured in different scenarios. Seven test protocols, which identify different training and testing images, are defined in [19] to evaluate verification algorithms. We selected protocol P in our experiments since it is the most challenging. The protocol specifies the partitioning of the database into two disjoint sets: a *development set* (26 subjects) and an *evaluation set* (26 subjects). For each set, 5 images from each person captured in the 1st session (Controlled scenario) are used as training images, while 2730 selected images captured in all three scenario are used for testing. Each set thus consists of 130 training images, and test images consisting of 1170 client accesses and 1560 impostor accesses [19]. The same normalization process as used in the identification test was similarly applied here. Thus, face images are rotated, translated and scaled according to the position of the eyes. The images are cropped to standard size (48×48) and rotated so that the eyes are placed at fixed points. To reduce illumination variations, all of the images are shifted and scaled such that the mean values of all pixels equals zero, while the standard deviation equals one.

5.2.2. Performance metrics

Two error-related metrics (FAR and FRR) are used to measure performance of the verification system. The false acceptance rate, or FAR, is the measure of the likelihood that the biometric security system will incorrectly accept an access attempt by an unauthorized user. The false rejection rate, or FRR, is the measure of the likelihood that the biometric security system will incorrectly reject an access attempt by an authorized user. The values of FAR and FRR are computed using:

$$FAR = \frac{n_{ac}}{n_u}; \quad FRR = \frac{n_{re}}{n_a} \tag{20}$$

where n_a is the number of access attempts by an authorized user, n_u is the number of access attempts by an unauthorized user, n_{ac} is the number of acceptances for unauthorized users and n_{re} is the number of rejections for authorized users.

A system can be tuned for a particular application by varying the value of these two metrics. A low value for both metrics is often desirable though, unfortunately, trying to minimise FAR or FRR requires a trade-off between the two metrics. Thus, system design requires the knowledge of the target application domain in order to determine the optimum balance between the two measures in practice. The Receiver Operating Curve (ROC) plots FAR versus FRR [17] for a system and can be used as a guide for the selection of an appropriate operating point for the system

Fig. 10. Example images in the BANCA database.

in relation to the target application. The closer the ROC-curve lies to the x and y axes, the lower the verification error and thus the more reliable the system. From the ROC-curve, the Equal Error Rate (*EER*) is defined as the point where the value of *FAR* equals the value of *FRR*. The value of *EER*, although not reflecting the differences in FAR/FRR trajectories, is often regarded as a useful composite system performance indicator for a verification system. Thus, from this point of view, the lower is the value of *EER*, the more reliable the system overall.

5.2.3. Parameter tuning and results on the development set

The first series of verification experiments were conducted using the *development set*. Using the extracted Gabor features and GDA subspace projection techniques for face verification, all parameters (e.g. feature dimension, distance measure and decision threshold, etc.) are optimized for the best performance. The ROC curves generated are displayed in Fig. 11, while Table 2 shows the *EER* and corresponding *FAR* and *FRR* of the Gabor feature based verification algorithms using different subspace projection techniques, with the method of PCA and LDA as the baseline. Comparing Gabor + PCA and Gabor + LDA with the corresponding baseline, we observe a large margin of improvement when Gabor features are used. Eleven percent improvement has been achieved for PCA when Gabor features instead of original images are used for verification. The observation verified the effectiveness of Gabor features for classification. Although a number of papers available in the literature report the advantages of LDA over PCA, this experiment reveals the LDA algorithms to perform less well than PCA. However, several papers have also reported similar results (see, for example [32,33]) to ours. As indicated in [33], there is no guarantee that LDA will outperform PCA when a small (or non-representative) training set is used. In these experiments, the training set consists only of images captured in the Controlled scenario, while the test images were captured using different cameras (Degraded scenario) and in significantly different poses (Adverse scenario). However, as a kernel version of LDA, GDA still achieves the best performance.

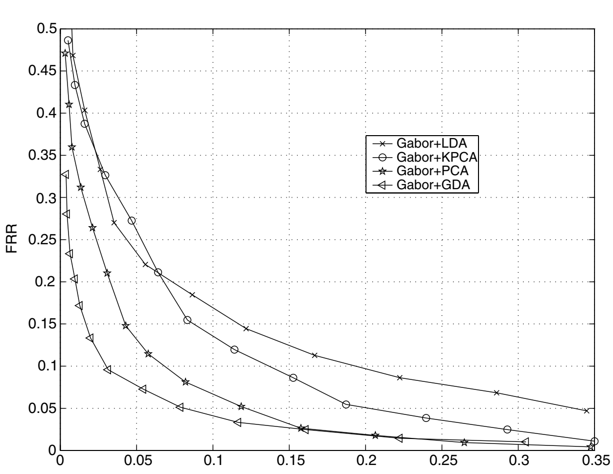

Fig. 11. ROC curve of Gabor based verification.

Table 2
Verification performance on the development set

Method	Distance measure	*FAR* (%)	*FRR* (%)	*EER* (%)
PCA	Ma	18.52	19.65	19.09
LDA	Nc	22.62	22.64	22.63
Gabor + PCA	Ma	8.20	8.11	8.15
Gabor + LDA	Nc	13.01	13.41	13.21
Gabor + KPCA	Ma	10.38	10.34	10.36
Gabor + GDA	Nc	6.02	5.89	5.96

The Mahalanobis distance measure (Ma) was found to be the most suitable measure for both PCA and KPCA, while the Nc distance measure is seen to be optimal for LDA and GDA. This observation is the same as that which was made in relation to the identification experiments. It can be seen from Table 2 that the Gabor + GDA method performs the best of the approaches considered, with only a 5.96% *EER*. The *EER* for the Gabor + KPCA method is around 10.36%, which is worse than the Gabor + PCA approach.

5.2.4. Generalization test using the tuned parameters

An independent *evaluation set* was designed in *protocol P* to test the generalization ability of verification algorithms. The *evaluation set* consists of the same number of subjects and images as that of the *development set*. However, the subjects of the *evaluation set are* distinct from those in the *development set*. With parameters adjusted and performance optimized using the *development set*, the generalization ability of algorithms can be further analyzed using the *evaluation set*. The value of *EER* attained with respect to the methods Gabor + PCA, Gabor + LDA, Gabor + KPCA and Gabor + GDA on the *evaluation set* are tabulated in Table 3. In each case, the system has been tuned to the *development set* in the first series of experiments, i.e., the decision threshold has been decided upon during the development phase. Again, the Gabor + GDA method achieves the lowest *EER*, while Gabor + PCA is better than Gabor + KPCA and Gabor + LDA.

5.2.5. Comparison with other methods

The 2004 international face verification competition [20], held in conjunction with the International Conference on Pattern Recognition (ICPR) 2004, has been adopted as an informal benchmark for comparisons within the research community. The competition adopted the BAN-

Table 3
Verification performance on the *evaluation set*

Method	*FAR* (%)	*FRR* (%)	*EER* (%)
Gabor + PCA	7.17	9.57	8.37
Gabor + LDA	7.56	10.51	9.04
Gabor + KPCA	8.26	13.33	10.79
Gabor + GDA	7.75	7.43	7.69

Table 4
Comparison with other verification methods

Method	*EER* (%)	
	Development set	Evaluation set
DCT-HMM	25.23	26.25
LDA	12.46	13.66
LDA + SVM	12.61	13.84
Elastic Bunch Graph Matching	12.10	16.80
Fusion of Gabor and global features	2.61	1.85
Gabor + GDA	5.96	7.69

CA database and protocol *P* for evaluation. Since the same database and protocol was used in our experiments, our verification results are directly comparable with those of the participants in the competition, providing a useful focus for further discussion. Table 4 shows the *EER* of a number of popular techniques on both the *development set* and *evaluation set*, which are directly extracted from the test report [20]. The results are very encouraging, with the *EER* returned by our method on both the *development set* and *evaluation set* seen to be significantly better than many popular methods reported in the literature, such as LDA, Elastic Bunch Graph Matching, HMM, etc. Since LDA can be implemented in various ways, the difference of the results for LDA reported in Tables 2 and 4 could be caused by the difference in our implementations and other researcher's. The method fusing Gabor with other features outperformed our method, suggesting that the performance of Gabor + GDA could be further improved by fusing other complementary features. However, such fusion could bring large computation burden to the system and make it not practical for real applications.

6. Discussion

In this paper, we have proposed a novel Gabor + GDA method for face identification and verification. The design of Gabor filters for facial feature extraction has also been discussed and experimentally tested for face recognition purposes. Extensive experiments have been conducted to evaluate the performance of the algorithm and compare it with alternative approaches, using the FERET and BANCA databases for reference. The proposed method achieves 97.5% recognition rate on the FERET database, which is significantly better than the Eigenface, Fisherface and HMM approaches. The method has also been extensively tested using the BANCA database, where the *development set* was used for algorithm optimisation and the *evaluation set* was used for generalization test. Our method outperformed all entries listed in Table 4 to the 2004 International Face Verification Competition, except one that fuses Gabor and global image features. However, the fusion scheme may bring large computational burden to the system.

Our verification tests support the claim by a number of other researchers [32,33] that PCA outperforms LDA when the training set is not representative. This suggests that the application itself should be considered when choosing recognition algorithms. The comparison among different distance measures such as the Euclidean, Mahalanobis and Correlation measures, clearly shows that PCA and KPCA favour the Mahalanobis distance measure, while LDA and GDA are better served by the Correlation measure. Since the Euclidean distance treats each component of the feature vector equally, the results suggest that more sophisticated metrics, (such as the Mahalanobis and Correlation measures) should be used for subspace-based methods. As a weighted distance metric, Mahalanobis weights each component with its variance. When PCA subspace is used, each component of the feature vector has been decorrelated with others. The motivations for adopting the Mahalanobis metric (i.e., the larger the variance, the more significant the associated component), are thus most exploited in the PCA space. With an aim to maximize the class separability, the components of the LDA feature vector have already been associated with their significance for classification. Inaccurate weighting of the Mahalanobis scheme could counteract such significance measures and direct correlation thus become more desirable.

Though the robustness and accuracy of Gabor + GDA has been extensively tested and evaluated we are currently developing methods to reduce the feature dimension further, using high level feature selection schemes such as boosting [34,35], mutual information [36], etc.

References

[1] M. Lades, J.C. Vorbruggen, J. Buhmann, J. Lange, C. Von der Malsburg, R.P. Wurtz, W. Konen, Distortion invariant object recognition in the Dynamic Link Architecture, IEEE Transactions on Computers 42 (3) (1993) 300–311.

[2] L. Wiskott, J.M. Fellous, N. Kruger, C. Von der Malsburg, Face recognition by elastic bunch graph matching, IEEE Transactions on Pattern Analysis and Machine Intelligence 19 (7) (1997) 775–779.

[3] A. Nefian, M. Hayes, An embedded hmm-based approach for face detection and recognition, in: Proceedings of IEEE International Conference On Acoustics, Speech and Signal Processing, 1999, pp. 3553–3556.

[4] M. Turk, A. Pentland, Eigenfaces for recognition, Journal of Cognitive Neuroscience 3 (1) (1991) 71–86.

[5] P.N. Belhumeur, J.P. Hespanha, D.J. Kriegman, Eigenfaces vs. Fisherfaces: recognition using class specific linear projection, IEEE Transactions on Pattern Analysis and Machine Intelligence 19 (7) (1997) 711–720.

[6] M.S. Bartlett, J.R. Movellan, T.J. Sejnowski, Face recognition by independent component analysis, IEEE Transactions on Neural Networks 13 (6) (2002) 1450–1464.

[7] B.A. Draper, K. Baek, M.S. Bartlett, J.R. Beveridge, Recognizing faces with PCA and ICA, Computer Vision and Image Understanding 91 (1–2) (2003) 115–137.

[8] P.J. Phillips, H. Moon, S.A. Rizvi, P.J. Rauss, The FERET evaluation methodology for face-recognition algorithms, IEEE Transactions on Pattern Analysis and Machine Intelligence 22 (10) (2000) 1090–1104.

[9] B. Scholkopf, S. Mika, C.J.C. Burges, P. Knirsch, K.R. Muller, G. Ratsch, A.J. Smola, Input space versus feature space in kernel-based methods, IEEE Transactions on Neural Networks 10 (5) (1999) 1000–1017.

[10] K.I. Kim, K. Jung, H.J. Kim, Face recognition using kernel principal component analysis, IEEE Signal Processing Letters 9 (2) (2002) 40–42.

[11] C.J. Liu, Gabor-based kernel PCA with fractional power polynomial models for face recognition, IEEE Transactions on Pattern Analysis and Machine Intelligence 26 (5) (2004) 572–581.

[12] M. Yang, Kernel Eigenfaces vs. Kernel Fisherfaces: face recognition using Kernel methods, in: Proceedings of the Fifth IEEE International Conference on Automatic Face and Gesture Recognition Washington, DC, 2002, pp. 205–211.

[13] Q.S. Liu, R. Huang, H.Q. Lu, S.D. Ma, Kernel-based nonlinear discriminant analysis for face recognition, Journal of Computer Science and Technology 18 (6) (2003) 788–795.

[14] G Baudat, F.E. Anouar, Generalized discriminant analysis using a kernel approach, Neural Computation 12 (10) (2000) 2385–2404.

[15] Cristianini Nello, Shawe-Taylor John, An introduction to Support Vector Machines and Other Kernel-based Learning Methods, Cambridge University Press, 2000.

[16] P.J. Phillips, Support vector machines applied to face recognition, in: Proceedings of the 1998 Conference on Advances in Neural Information Processing Systems II, MIT Press, 1999, pp. 803–809.

[17] K. Jonsson, J. Kittler, Y.P. Li, J. Matas, Support vector machines for face authentication, Image and Vision Computing 20 (5–6) (2002) 369–375.

[18] C.J. Liu, H. Wechsler, Gabor feature based classification using the enhanced Fisher linear discriminant model for face recognition, IEEE Transactions on Image Processing 11 (4) (2002) 467–476.

[19] E.B. Baillere, S. Bengio, The BANCA database and evaluation protocol, in: Proceedings of the 4th International Conference on Audio- and Video-Based Biometric Person Authentication Guilford, UK, 2003, pp. 625–638.

[20] K. Messer, J. Kittler, M. Sadeghi, M. Hamouz, A. Kostin, F. Cardinaux, S. Marcel, S. Bengio, C. Sanderson, N. Poh, Y. Rondriguez, J. Czyz, L. Vandendorpe, C. McCool, S. Lowther, S. Sridharan, V. Chandran, R.P. Palacios, E. Vidal, L. Bai, L. Shen, Y. Wang, Y.H. Chiang, H.C. Liu, Y.P. Huang, A. Heinrichs, M. Miiller, A. Tewes, C.v.d. Malsburg, R. Wiirtz, Z.G. Wang, F. Xue, Y. Ma, Q. Yang, C. Fang, X.Q. Ding, S. Lucey, R. Goss, H. Schneiderman, Face authentication test on the BANCA database, in: Proceedings of International Conference on Pattern Recognition, Cambridge, UK, 2004, pp. 523–532.

[21] A.K. Jain, F. Farrokhnia, Unsupervised Texture Segmentation Using Gabor Filters, Pattern Recognition 24 (12) (1991) 1167–1186.

[22] T.P. Weldon, W.E. Higgins, D.F. Dunn, Efficient Gabor filter design for texture segmentation, Pattern Recognition 29 (12) (1996) 2005–2015.

[23] Y. Hamamoto, S. Uchimura, M. Watanabe, T. Yasuda, Y. Mitani, S. Tomita, A Gabor filter-based method for recognizing handwritten numerals, Pattern Recognition 31 (4) (1998) 395–400.

[24] C.J. Lee, S.D. Wang, Fingerprint feature extraction using Gabor filters, Electronics Letters 35 (4) (1999) 288–290.

[25] K. Fukunnaga, Introduction to Statistical Pattern Recognition, second ed., Academic Press, 1991.

[26] D.L. Swets, J.J. Weng, Using discriminant eigenfeatures for image retrieval, IEEE Transactions on Pattern Analysis and Machine Intelligence 18 (8) (1996) 831–836.

[27] J.W. Lu, K.N. Plataniotis, A.N. Venetsanopoulos, Regularized discriminant analysis for the small sample size problem in face recognition, Pattern Recognition Letters 24 (16) (2003) 3079–3087.

[28] C.J. Liu, H. Wechsler, Enhanced fisher linear discriminant models for face recognition, in: Proceedings of International Conference on Pattern Recognition, 1998, pp. 1368–1372.

[29] H. Yu, H. Yang, A direct LDA algorithm for high-dimensional data – with application to face recognition, Pattern Recognition 34 (10) (2001) 2067–2070.

[30] M.J. Er, S.Q. Wu, J.W. Lu, H.L. Ton, Face recognition with radial basis function (RBF) neural networks, IEEE Transactions on Neural Networks 13 (3) (2002) 697–710.

[31] L. Bai, L. Shen, Combining wavelet and HMM for face recognition, in: Proceedings of the 23rd Artificial Intelligence Conference, Cambridge,UK, 2003, pp. 227–234.

[32] J.R. Beveridge, K. She, B. Draper, A nonparametric statistical comparison of principal component analysis and linear discriminant subspaces for face recognition, in: Proceeding of IEEE International Conference on Computer Vision and Pattern Recognition, 2001.

[33] A.M. Martinez, A.C. Kak, PCA versus LDA, IEEE Transactions on Pattern Analysis and Machine Intelligence 23 (2) (2001) 228–233.

[34] P. Yang, S.G. Shan, W. Gao, S.Z. Li, D. Zhang, Face recognition using ada-boosted gabor features, in: Sixth IEEE International Conference on Automatic Face and Gesture Recognition, Proceedings, 2004, pp. 356–361.

[35] L. Shen, L. Bai, Mutual Boost learning for selecting Gabor features for face recognition, Pattern Recognition Letters, in press, doi:10.1016/j.patrec.2006.02.005.

[36] L. Shen, L. Bai, Information theory for Gabor feature selection for face recognition, Eurasip Journal on Applied Signal Processing, in press, doi:10.1155/ASP/2006/30274.