PERGAMON

Pattern Recognition 36 (2003) 313–328

PATTERN RECOGNITION
THE JOURNAL OF THE PATTERN RECOGNITION SOCIETY

www.elsevier.com/locate/patcog

Face authentication for multiple subjects using eigenflow

Xiaoming Liu, Tsuhan Chen*, B.V.K. Vijaya Kumar

Department of Electrical and Computer Engineering, Carnegie Mellon University, 5000 Forbes Ave., Pittsburgh, PA 15213-3890, USA

Received 21 December 2001

Abstract

In this paper, we present a novel scheme for face authentication. To deal with variations, such as facial expressions and registration errors, with which traditional intensity-based methods do not perform well, we propose the eigenflow approach. In this approach, the optical flow and the optical flow residue between a test image and an image in the training set are first computed. The optical flow is then fitted to a model that is pre-trained by applying principal component analysis to optical flows resulting from facial expressions and registration errors for the subject. The eigenflow residue, optimally combined with the optical flow residue using linear discriminant analysis, determines the authenticity of the test image. An individual modeling method and a common modeling method are described. We also present a method to optimally choose the threshold for each subject for a multiple-subject authentication system. Experimental results show that the proposed scheme outperforms the traditional methods in the presence of facial expression variations and registration errors.

Keywords: Eigenflow; Individual eigenspace; Face authentication; Principal component analysis; Optical flow

1. Introduction

For decades human face recognition has drawn considerable interest and attention from many researchers [1]. A general statement of this problem can be formulated as follows. Given still or video images of a scene, identify one or more persons in the scene using a stored database of faces [2]. A system that performs face recognition will have many applications such as nonintrusive identification and authentication for credit cards usage, nonintrusive access control to buildings, and identification for law enforcement.

Face authentication [3] is a research field related to face recognition. The difference between face recognition and face authentication is that, in the former, the system has to determine the identity of the subject, while in the latter, the system needs to verify the claimed identity of the subject. Usually, similar algorithms can be used for both recognition and authentication. However, there are some major differences between these two cases. For example, the recognition application is usually computationally intensive compared to the authentication application. Also, performance evaluation methods are different, which we will elaborate on later. Face recognition and authentication are challenging as the human face can always undergo significant variations in appearance because of changes in facial expressions, poses, scales, shifts, and lighting conditions.

In this paper, we propose a new approach to performing face authentication that is tolerant to facial expression variations and registration errors. Optical flow is used to capture face appearance motion when there are variations in facial expressions. For example, the optical flow between the neutral and happy expressions of one subject tells us how this subject smiles. After we apply principal component analysis (PCA) to these optical flows, we obtain an eigenspace spanned by its eigenvectors. This eigenspace models all possible expression variations. We call this the *eigenflow* approach in this paper. Similarly, optical flow and eigenflow

* Corresponding author. Tel.: +1-412-268-7536; fax: +1-412-268-3890.

E-mail addresses: xiaoming@andrew.cmu.edu (Xiaoming Liu), tsuhan@cmu.edu (T. Chen), kumar@ece.cmu.edu (B.V.K. Vijaya Kumar).

0031-3203/02/$22.00
PII: S0031-3203(02)00033-X

can also be used to model other variations, such as registration errors, including shifts, scales, and rotations.

Given a test image, the optical flow and the optical flow residue between a test image and an image in the training set are computed. The optical flow is fitted to the eigenflow model. The eigenflow residue is optimally combined with the optical flow residue using linear discriminant analysis (LDA), and the result is used to determine the authenticity of the test image.

The advantage of this proposed approach is in its tolerance to expression variations and registration errors. Because the traditional PCA approach does not deal well with registration errors [4], it requires that all the training and test images to be precisely registered, sometimes manually. Our eigenflow-based approach, however, is tolerant to such registration errors; even face images subjected to significant registration errors can be authenticated correctly.

1.1. Previous work

Comprehensive surveys of human and machine recognition techniques can be found in Refs. [2,5,6]. There are mainly two kinds of face recognition systems: one is based on the feature matching; the other is based on the template matching. In the latter, applying PCA in the pixel domain (also known as the eigenface approach [7]) plays a fundamental role. It has the advantage of fast computation, stable performance for the case of frontal face recognition with reasonable constraints on illumination, expression variations, etc. Turk and Pentland [7] used the whole training image set to train one eigenspace, which combines the intra- and inter-subject variations into a single model. In this paper, we propose classification metrics based on the individual eigenspaces, because they can model the intra- and inter-variations separately, which is much more appropriate for classification.

The eigenface approach has been revised to dealing with face image variability [8]. However, both the original PCA and their revised versions are intensity-based methods. Therefore, they cannot model variations such as shifts, rotations, scales, expressions or lighting variations.

Optical flow methods are generally used for motion analysis. Using two or more consecutive frames of an image sequence, a two-dimensional vector field, called the optical flow, is computed to estimate the most likely displacement of image pixels from one frame to another. Some researchers have used optical flow in the analysis of human expressions for the purpose of expression recognition [9,10]. Also Kruizinga and Petkov [11] proposed to utilize optical flows in person identification. However, they only considered the optical flow residue as the criterion of classification, while we propose to make use of the eigenflow residue, which appears to exhibit better classification ability than the former.

Essentially, optical flow can provide us the visual motion information about face images. Moghaddam et al. also proposed modeling visual motion in Ref. [12]. They determined pixel difference between images, and utilized the Bayesian approach to modeling the pixel difference for all the subjects. In our case, first the optical flow is used to obtain the motion field between images and then PCA is applied to model facial motion for individual subjects.

Based on our approach, we perform experiments on different face databases. Better performance has been obtained compared to the traditional PCA approach, especially when the test face has expression variations and registration errors.

1.2. Paper outline

In Section 2, we propose an improved solution for face authentication compared to the traditional PCA approach. Since an individual eigenspace models the intra-subject variations well, the distance to this space can be thought as one measurement of the inter-subject variations. Both the traditional and our individual PCA approach will be presented in detail.

In Section 3, we present the individual eigenflow-based approach in detail. One eigenflow space is trained using optical flows between all the images of one subject. Two residues, the eigenflow residue and the optical flow residue, are combined by LDA for authentication.

In practical authentication applications, it is possible that the subject may not provide training images containing expression variations. How can face images exhibiting expression variations still be authenticated based on the neutral training faces? We discuss this problem in Section 4 and suggest a solution—a common eigenflow space is trained before it is delivered to any specific applications. This space describes all the possible expression variations that might appear in any subjects.

In Section 5, we introduce a new approach to optimizing and evaluating the authentication performance for multiple subjects. Different thresholds are customized for each subject. Then multiple receiver operating characteristic (ROC) curves from different subjects are combined into one curve, which optimizes the performance of the whole system.

Experiments based on different face data sets are presented in Section 6. Every experiment has its own training and test data set, and illustrates the usage of our new approach. Finally in Section 7, we provide our conclusions.

2. Universal and individual approaches to PCA

Turk and Pentland [7] introduced the eigenface approach to performing face recognition. To construct an eigenspace, face images from all training subjects are used. We call the resulting eigenspace a universal eigenspace because it represents the appearance variations among all the subjects. Within this space, the face images of a single subject typically cluster around a certain region. For face authentication, we may be able to improve the performance by using eigenspaces tuned to each individual. In this section, we first describe the algorithm based on the universal eigenspace. Then in Section 2.2 we introduce a new approach, the individual PCA approach.

2.1. Universal PCA approach

Suppose there are M training face images for each of the K subjects. Let each face image $I(x, y)$ be a two-dimensional $N \times N$ array of pixel values. An image may also be represented (after scanning) as a vector of dimension N^2, where one image corresponds to one point in the N^2-dimensional image space. Let us denote each face image of the training set as $\mathbf{f}_{ij}$, a $N^2 \times 1$ vector, where i and j denote the subject index and the face image index, respectively, and $0 \leqslant i \leqslant K-1, 0 \leqslant j \leqslant M-1$. The average face vector $\mathbf{g}$ is defined by

$$\mathbf{g} = \frac{1}{M \times K} \sum_{i=0}^{K-1} \sum_{j=0}^{M-1} \mathbf{f}_{ij}. \tag{1}$$

The difference between each training face and the average is denoted by the vector $\mathbf{s}_{ij} = \mathbf{f}_{ij} - \mathbf{g}$. These difference vectors form a $N^2 \times MK$ matrix, $\mathbf{A} = [\mathbf{s}_{00}, \mathbf{s}_{01}, \ldots, \mathbf{s}_{K-1,M-1}]$. We apply PCA to these difference vectors by finding a set of Q orthonormal eigenvectors, $\mathbf{u}_n$, corresponding to the largest eigenvalues of matrix $\mathbf{AA}^{\mathrm{T}}$, i.e.,

$$\mathbf{AA}^{\mathrm{T}}\mathbf{u}_n = \lambda_n \mathbf{u}_n \quad n = 0, 1, \ldots, Q-1, \tag{2}$$

where $\lambda_0, \lambda_1, \ldots, \lambda_{Q-1}$ are nonnegative and in a decreasing order. However, the matrix $\mathbf{AA}^{\mathrm{T}}$ is $N^2 \times N^2$, and determining N^2 eigenvectors can be computationally intensive. Usually the number of training faces, $M \times K$, is much smaller than N^2. So we first determine the eigenvectors, $\mathbf{u}_n$, of a $MK \times MK$ matrix $\mathbf{A}^{\mathrm{T}}\mathbf{A}$ i.e.,

$$\mathbf{A}^{\mathrm{T}}\mathbf{A}\mathbf{u}_n = \lambda_n \mathbf{u}_n \tag{3}$$

Pre-multiplying Eq. (3) by $\mathbf{A}$ and comparing to Eq. (2), we can see that $\mathbf{u}_n = \mathbf{A}\mathbf{u}_n \lambda_n^{-1/2}$. These eigenvectors form an orthonormal basis set of a new feature space, called the eigenspace. Essentially, it is a subspace representation of all the faces. Thus, we can transform each face image, $\mathbf{f}_{ij}$, from the image space to the eigenspace as follows:

$$w_n = \mathbf{u}_n^{\mathrm{T}}(\mathbf{f}_{ij} - \mathbf{g}), \quad n = 0, 1, \ldots, Q-1. \tag{4}$$

Since each face image can be described as a vector $\mathbf{p}_{ij} = [w_0, w_1, \ldots, w_{Q-1}]^{\mathrm{T}}$ in the eigenspace, the Mahalanobis distance [13] can be used to perform face authentication. After all training faces, $\mathbf{f}_{ij}$, are projected into the eigenspace, the mean vector and the covariance matrix corresponding to each training subject are determined via $\mathbf{p}_{ij}$,

$$\mathbf{m}_i = \frac{1}{M} \sum_{j=0}^{M-1} \mathbf{p}_{ij}, \tag{5}$$

$$\mathbf{C}_i = \frac{1}{M} \sum_{j=0}^{M-1} (\mathbf{p}_{ij} - \mathbf{m}_i)(\mathbf{p}_{ij} - \mathbf{m}_i)^{\mathrm{T}}. \tag{6}$$

The training faces of K subjects in the eigenspace represent K classes, with $\mathbf{m}_i$ and $\mathbf{C}_i$ describing the mean vector and the covariance matrix of Class i, respectively.

The above estimation of the mean and the covariance matrix is carried out in the training stage. During the test stage of this authentication system, suppose one face, $\mathbf{f}$, is presented to the system, and is claimed to be Subject i. First, $\mathbf{f}$ is projected to the ith eigenspace as in Eq. (4) to obtain the projected vector $\mathbf{p}$. Then the following Mahalanobis distance is used to measure the distance between the test face $\mathbf{f}$ and the ith class.

$$d = (\mathbf{p} - \mathbf{m}_i)^{\mathrm{T}} \mathbf{C}_i^{-1} (\mathbf{p} - \mathbf{m}_i). \tag{7}$$

Finally by comparing d with a pre-selected threshold, the test face, $\mathbf{f}$, can be accepted as from Subject i or rejected as from imposters. The advantage of using the Mahalanobis distance is that it takes into account the class covariance in determining the distance to that class.

The eigenvector determination is computationally expensive, especially when the number of training images is large. The power method [14] is one approach to efficiently determining the dominant eigenvectors. Instead of determining all the eigenvectors, the power method obtains only the dominant eigenvectors, i.e., eigenvectors associated with the largest eigenvalues. We also used an eigenspace update algorithm [15] with the benefit of incrementally updating the eigenvectors based on the new training image.

2.2. Individual PCA approach

For the universal eigenspace, it represents not only the personal identity, the inter-variations between different training subjects, but also the intra-variations of each subject, such as due to expression changes, illumination variability, age, etc. However, what we need for the authentication is robustness to expressions and illumination variations within a single subject. This observation suggests one potential metric for face authentication. The residue of a test vector to that vector's individual eigenspace (i.e., the squared norm of the difference between a test vector and its representation in the eigenspace) is a good measure for authentication.

Throughout the rest of this paper, we will focus for clarity on one specified subject. So for a data set of K subjects, the same approach can be repeated K times. In the individual PCA approach, one eigenspace is constructed for each training subject. As the following equation indicates, the average face will be different for each subject:

$$\mathbf{g}_i = \frac{1}{M} \sum_{j=0}^{M-1} \mathbf{f}_{ij}. \tag{8}$$

Now each face differs from the average by $\mathbf{s}_{ij} = \mathbf{f}_{ij} - \mathbf{g}_i$. Also $\mathbf{A}_i = [\mathbf{s}_{i,0}, \mathbf{s}_{i,1}, \ldots, \mathbf{s}_{i,M-1}]$ is different for each subject. By using of the same algorithm as the last section, we can train an individual eigenspace for each subject.

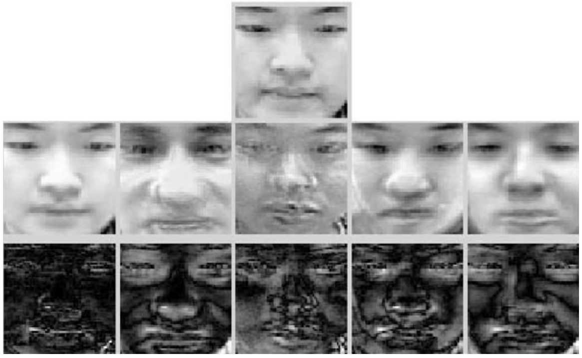

Fig. 1. One sample face, five reconstructed faces and five residue images.

Given a test image $\mathbf{f}$ with its claimed identity, it is projected to the eigenspace of the claimed subject as follows:

$$w_n = \mathbf{u}_{i,n}^{\mathrm{T}}(\mathbf{f} - \mathbf{g}_i) \quad \text{if } \mathbf{f} \textit{ claims to be Subject } i, \tag{9}$$

where $\mathbf{u}_{i,n}$, is the eigenvector of the eigenspace. From the projected vector, the face image can be reconstructed from

$$\hat{\mathbf{s}} = \sum_{n=0}^{Q-1} W_n \mathbf{u}_{i,n} \tag{10}$$

by keeping only the eigenvectors corresponding to the Q largest eigenvalues. Since the eigenspace is only an idealized representation for one subject, it cannot represent all manifestations of that subject's face perfectly, i.e., there will be a residue (i.e., the squared error) between the face image and its reconstructed version. Such residue can be seen clearly in Fig. 1. The face in the first row is a sample face to be authenticated. We show in the second row the reconstructed faces in five individual eigenspaces (corresponding to five different subjects). It is clear that the left-most image in the second row is the most similar to the test image. The residue images between the test face and its corresponding reconstructed images for the five subjects are shown in the third row. We can see that the residue image has low intensity only when the sample face is projected to its own space. Thus, the low energy in the residue image can be a good criterion to authenticate faces.

The residue is defined as the squared distance between the mean-adjusted test input image $\mathbf{s} = \mathbf{f} - \mathbf{g}_i$ and the reconstructed image $\hat{\mathbf{s}}$, i.e.,

$$e = \|\mathbf{s} - \hat{\mathbf{s}}\|^2. \tag{11}$$

This is similar to the residue defined in the subspace method [16]. There is one method to speed up the computation of this residue. Instead of first determining the reconstructed image $\hat{\mathbf{s}}$ and then using Eq. (11), we can use the following to determine e:

$$e = \|\mathbf{s}\|^2 - \sum_{n=0}^{Q-1} w_n^2. \tag{12}$$

In the next section, we will extend PCA from the pixel domain to the optical flow domain, and the residue will be called the *eigenflow residue*.

3. Individual eigenflow-based face authentication

The traditional PCA approach is not as robust as needed to expression variations, shift, rotation, and scale changes. Because PCA is an intensity-based approach, its authentication performance will degrade quickly when the appearance of a subject's face changes significantly, which occurs in the presence of expression changes and registration errors. In this section, we propose a new method based on optical flow to deal with such variations in face images.

3.1. Optical flow for face images

Essentially optical flow [17] is an approximation of the velocity field. It characterizes approximately the motion of each pixel between two images.

If two face images, which show different expressions of the same subject, are fed into the optical flow algorithm, the resultant motion field will emphasize the regions of facial features, such as eyes and the mouth. This is illustrated in Fig. 2. The left half of the figure shows two face images from the same subject, but with different expressions. The resulting optical flow is shown below these figures. By using the first image and the optical flow, we can construct a predicted image that is close to the second image. This is similar to video coding where we can use the previous frame with motion vectors to form a prediction frame that is close the current frame. The third figure in the top row is the difference between the prediction obtained via the optical flow and the second image. We call it an *optical flow residue* image. For the same subject, this residue image would have low energy because the motion of most pixels can be well modeled by the optical flow. The second set shows the same figure except that the two input images are from two different subjects. Obviously, the optical flow looks more irregular in this case. Also the residue image of motion compensation has more "error". These two clues can help discriminating these two cases, which is the task of face authentication.

The same idea can be applied to images with registration errors. Because the traditional PCA approach is unacceptably sensitive to registration errors, even small shifts in input images can make the system performance degrade significantly. However, face images are usually difficult to register precisely, especially in a live authentication system. Therefore, we want to use the optical flow to build a system that is tolerant to different kinds of registration errors. In Fig. 3, the second image in the left column is an up-shifted version of the first image. The optical flow shown below captures most of its motion around facial features, and also the residue image has relatively small intensity. The right column shows images of different subjects leading to an

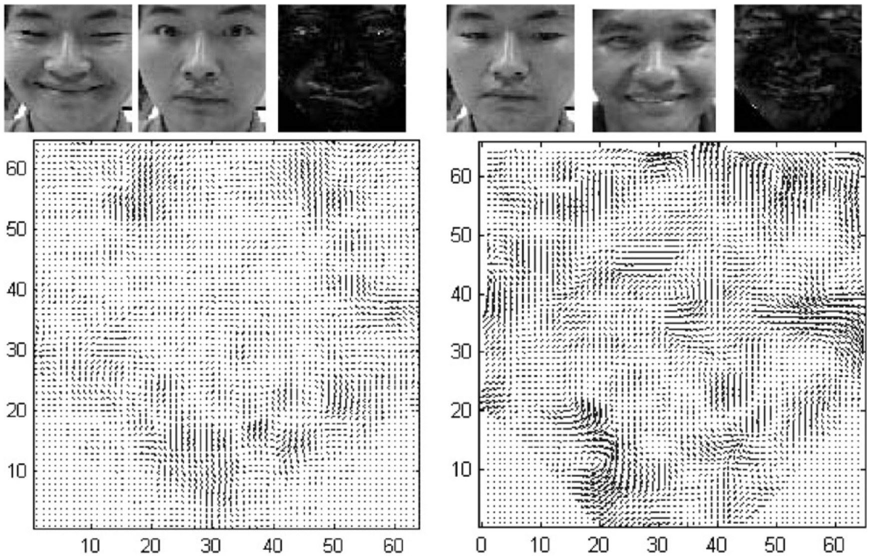

Fig. 2. Applying optical flow on images with different expressions.

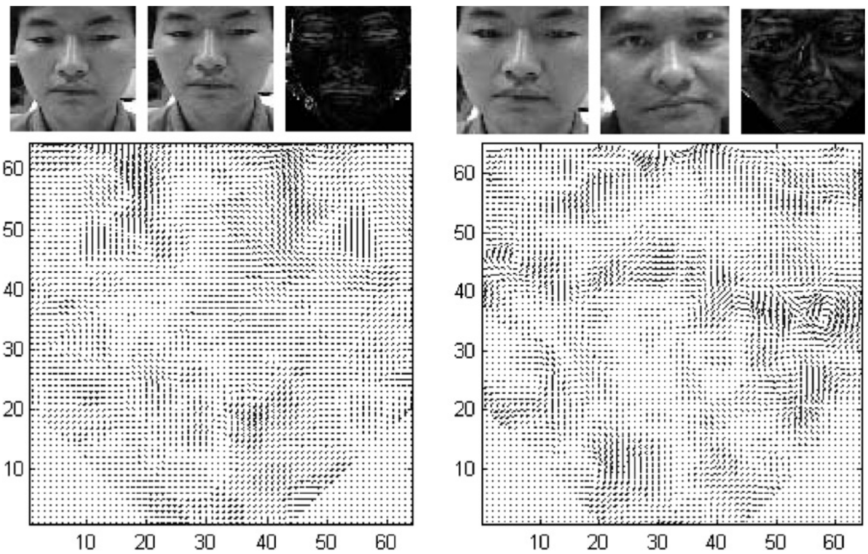

Fig. 3. Applying optical flow on images with registration errors.

optical flow that appears to be random, and the residue image has larger intensity.

3.2. Eigenflows

Since the optical flow provides a useful pattern for classifying personal identity, we propose to use PCA to model this pattern.

Given two face images, there are some recommended preprocessing steps prior to the optical flow determination. Since some regions of face images will always contain the background, it is better to crop the image before the optical flow computation. We also consider down-sampling the optical flow in order to clean up the noise motion vector and speed up the PCA process.

Following the approach described in Section 2.2, optical flow vectors are regarded as sample vectors for training. Suppose that in the training data set, there are a few images with different expressions for each subject, such as five images shown in Fig. 4. Using these images, twenty optical flow images (corresponding to twenty pairs) can be obtained through the process in Fig. 13. The three

Fig. 4. Five expression images used for training eigenflows.

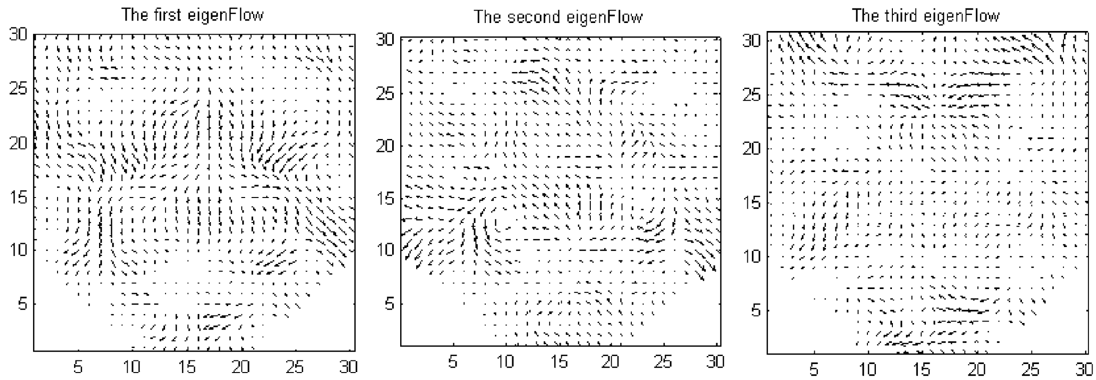

Fig. 5. The first three eigenflows trained from expression images of one subject. Some prominent movements of facial features, such as mouth corners, eyebrows, nasolabial furrows, can be seen from them.

principal eigenflows of twenty optical flow images are shown in Fig. 5. Obviously, large motion can be observed in the region of facial features, such as mouth corners, eyebrows and nasolabial furrows. So all the expression variations occurring in a single subject can be represented by a space spanned by these eigenflows. In contrast, the optical flow between this subject and other subjects cannot be represented well by this space, which results in a large residue. We call this the *eigenflow residue*. Similar to Section 2.2, different subjects would result in a relatively large eigenflow residue. Thus, the eigenflow residue can be a useful feature for authentication.

Similarly, eigenflows can be used to model the optical flow caused by image registration errors. In our approach, we can either synthesize images with registration errors from one well-registered image, or collect images with errors in the live application. By perturbing the face cropping operation (to be explained in Section 6.1), we synthesize images with registration errors. Some of the synthesized images are shown in Fig. 6. We can see that different kinds of registration errors, such as shifts, rotations, scales, appear in this set. Eigenflows based on them are shown in Fig. 7. Again the motion vectors in eigenflows indicate the actual motion patterns appearing in the training set. Comparing Fig. 7 with Fig. 5, we see that most motion in the former is inside the face region, such as mouth corners and eyebrows, while in the latter, the global motion is more dominant, for example, shifting to the left-up and right-up corner.

3.3. Linear discriminant analysis (LDA)

LDA is used to derive low-dimensional features from a high-dimensional feature space. Essentially, LDA is a feature space transformation and can be applied to two or more classes.

In face authentication, we need a good feature space, using which the subject can be clearly discriminated from imposters. So face authentication is equivalent to a 2-class classification problem, while one class being the *self* class, the other being the *imposter* class. When a test image comes in, we need to classify it into one of these two classes.

From the above description, both the eigenflow residue and the optical flow residue have the ability to discriminate between two classes. We can combine the discrimination ability of these two residues using LDA. Each optical flow image can be represented as one point in a two-dimensional feature space where one axis represents the eigenflow residue and the other represents the optical flow residue. One such space for a subject is shown in Fig. 8. Now, we utilize the LDA [13] approach to finding an optimal one-dimensional feature space, which can separate these two classes.

First, all the training optical flow images from Subject 0 are projected to this space, and they are regarded as the self class. Also we project the optical flow between Subject 0 and other $K-1$ subjects to the eigenflow space of Subject 0, and regard them as the imposter class. Now the mean and

Fig. 6. Synthesized training images with registration errors.

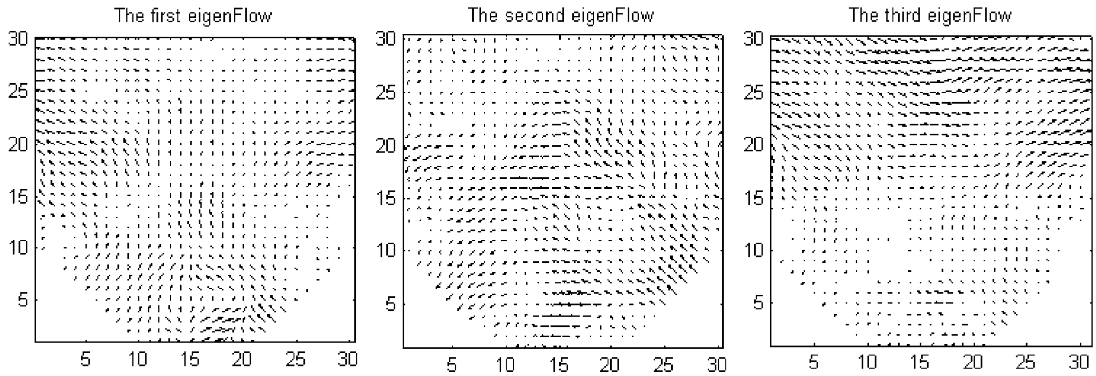

Fig. 7. The first three eigenflows trained from synthetic images of one subject. Some prominent motions, such as shifts, rotations, scales, can be seen from them.

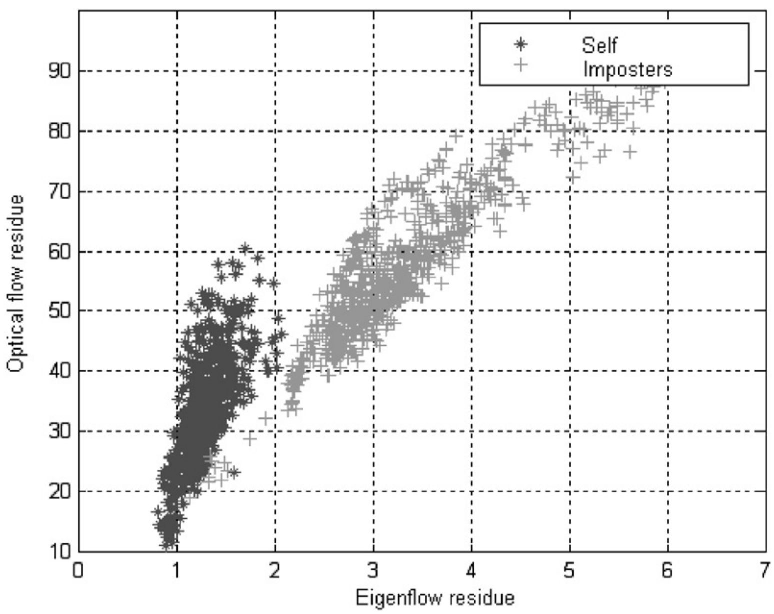

Fig. 8. The two residues shown in the feature space of one subject.

the covariance of these two classes can be estimated using Eqs. (5) and (6).

LDA is carried out via the scatter matrix analysis. For a 2-class problem, the within- and between-class scatter matrices $\mathbf{S}_w, \mathbf{S}_b$ are computed as follows:

$$\mathbf{S}_w = \tfrac{1}{2}(\mathbf{C}_s + \mathbf{C}_i), \tag{13}$$

$$\mathbf{S}_b = (\mathbf{m}_s - \mathbf{m}_i)(\mathbf{m}_s - \mathbf{m}_i)^{\mathrm{T}}, \tag{14}$$

where $\mathbf{m}_s, \mathbf{m}_i, \mathbf{C}_s, \mathbf{C}_i$ are the mean vectors, the covariance matrices of the self class, and the imposter class, respectively. Here $\mathbf{S}_w$ is the within-class scatter matrix showing the average covariance matrix of sample points of two classes in the two-dimensional space. Similarly $\mathbf{S}_b$ is the between-class scatter matrix, representing the scatter of two classes.

One measurement for quantifying the discriminatory power is the ratio of the determinant of the between-class scatter matrix to the within-class scatter matrix [19]:

$$J(\mathbf{w}) = \frac{|\mathbf{w}^{\mathrm{T}}\mathbf{S}_b\mathbf{w}|}{|\mathbf{w}^{\mathrm{T}}\mathbf{S}_w\mathbf{w}|}. \tag{15}$$

This is called the fisher ratio, which can be maximized by solving the generalized eigenvalue problem [20]:

$$\mathbf{S}_b\mathbf{w} = \mathbf{S}_w\mathbf{w}\lambda_w. \tag{16}$$

Since the rank of $\mathbf{S}_b$ is equal to 1 and $\mathbf{S}_w$ is invertible, the projection vector that maximizes $J(\mathbf{w})$ can be shown to be the following [13]:

$$\mathbf{w} = \frac{\mathbf{S}_w^{-1}(\mathbf{m}_s - \mathbf{m}_i)}{\|\mathbf{S}_w^{-1}(\mathbf{m}_s - \mathbf{m}_i)\|}. \tag{17}$$

Once we determine the vector $\mathbf{w}$, each sample in the original feature space can be transformed to a one-dimensional space, using which authentication can be carried out. From Fig. 8, it is clear that the eigenflow residue has a better discrimination ability for these two classes compared to the optical flow residue.

We have already shown the case of combining two sources of residue information using LDA. We also tried to combine information from two more features, such as the Mahalanobis distance based on the universal eigenspace (introduced in Section 2.1), and the residue in the individual eigenspace (introduced in Section 2.2). However, we found that adding those features does not improve the performance significantly. Therefore, we did not use them in authentication.

4. Common eigenflow-based face authentication

In the approaches described above, all the training images should exhibit expression variations if the test images are expected to have such variations. Suppose there are only neutral faces available for the training data set, but we still wish to have a face authentication system tolerant to expression variations. The traditional PCA approach will not work well in this case because of the significant appearance differences between faces with expressions and neutral faces. To deal with this difficulty, we propose the *common eigenflow*-based approach.

We observe that people exhibit similar facial motions for various expressions. Although individual details may vary slightly in changing expressions, the facial movements should be similar. For example, when people smile, usually the mouth opens wider and the cheek moves up. Based on this observation, the common eigenflow space is used to model possible expression variations occurring in all subjects. Then any expression changes within one subject can be well represented by this space, which should lead to a relatively small eigenflow residue. Also any expression changes between different subjects cannot be modeled by it, i.e., the eigenflow residue will be relatively larger. This is the key point in accepting or rejecting the identity claim. In this approach, there are two training stages: the eigenflow training and the LDA weights training.

Now let us discuss how to train this common eigenflow space in the first stage. First, we collect images with five kinds of expressions for many subjects. Then for each subject, optical flows can be determined between any two expressions of that subject. Using all these intra-subject optical flows, a common eigenflow space can be trained via PCA. The first three eigenflows are shown in Fig. 9.

Another thing that should be accomplished in the training stage is the training of LDA weights. In the test stage, LDA weights are used to combine the eigenflow residue and the optical flow residue. One choice is to train the LDA weights using the neutral training faces. However, because those training images are all neutral, the weights obtained from LDA training may not be suitable for the expression changes in the test set. To solve this problem, we propose an approach called *common weights.*

In the eigenflow training stage, for each subject we will get a 2-class plot in the feature space, such as in Fig. 8. Now in the training stage of common weights, we can plot the 2-class plots from all subjects together. Thus, one space is formed as in Fig. 10. The difference between Figs. 8 and 10 is that Fig. 8 shows the 2-class distribution for one specific subject, while Fig. 10 shows the 2-class distribution from all subjects. Applying LDA to this new space, common weights can be obtained. These common weights tell us how much each residue can contribute to the final measurement.

Now given any test image that claims to be Subject i, first the optical flow between the test image and the neutral training face of Subject i is calculated, then two residues are obtained, finally by using the common weights these residues can be combined into one value, which is used for authentication.

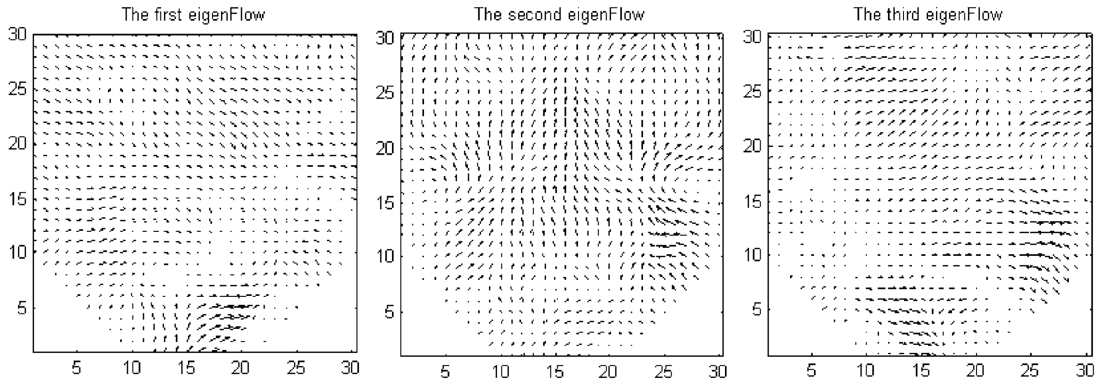

Fig. 9. Common eigenflows trained from optical flows between all expression variations of intra-subject.

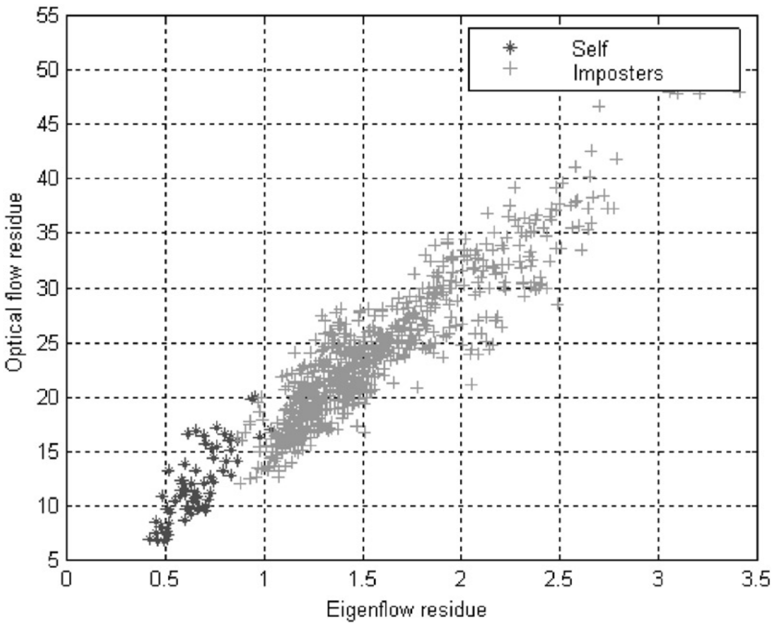

Fig. 10. The two classes of all subjects are plotted in one space. Common weights can be obtained by applying LDA to this space.

5. Evaluation of authentication for multiple users

One way to evaluate the performance of the authentication system is to plot the false accept rate (FAR) versus the false reject rate (FRR). If a subject claims to be himself, we call it a client claim. If a subject claims to be someone else, it is an imposter claim. FAR is the ratio of the false accepted number to the number of the imposter claims; FRR is the ratio of false rejected number to the number of client claims. Different pairs of FAR and FRR values can be computed by setting different threshold T. Therefore, we can create a plot of FRR versus FAR with T being an implicit parameter. Such a plot is called the ROC curve. It represents the performance of an authentication system. An ROC curve closer to the axes indicates better performance.

In the traditional face authentication system, usually there is only one threshold for all the subjects. However, we can customize thresholds for each subject. In this case, each subject will have its own individual ROC curve. Now the question is how we combine the multiple individual ROC curves into one curve, which characterizes the performance of the whole authentication system.

Given K ROCs, one for each subject, we can combine them into one ROC using the following algorithm. We describe each ROC by sampling False Accept points in the FAR coordinate. The sample point starts from FAR equal to

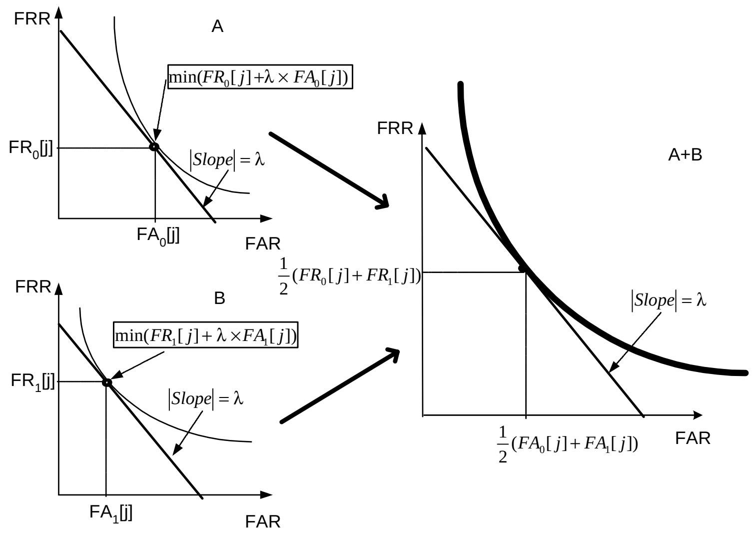

Fig. 11. Combining two ROC curves into one.

0, increases by d each time, and we do this n times. So FR_i and FA_i can uniquely determine the ith ROC curve. Since we want to find a set of thresholds to make the combined ROC close to the origin as much as possible, let us define the problem first.

Problem Statement.

Given K ROCs, each with $n+1$ sample points on the curve
$FR_i = \{a_{i0}, a_{i1}, \ldots, a_{ij}, \ldots, a_{in}\}, i = 0, 1, \ldots, K-1;$
$FA_i = \{0, d, \ldots, j \times d, \ldots, n \times d\}, i = 0, 1 \ldots, K-1;$
Find a set of thresholds, such that $\forall j$

$1/K \sum_{i=0}^{K-1} FA_i[j] \leqslant CFA[j]$, where $CFA[j]$ is predefined constants, and in the mean while $1/K \sum_{i=0}^{K-1} FR_i[j]$ is minimized.

This problem is equivalent to the following one,

$$\min \left(\frac{1}{K} \sum_{i=0}^{K-1} FR_i[j] + \lambda \times \frac{1}{K} \sum_{i=0}^{K-1} FA_i[j] \right)$$
$$= \frac{1}{K} \sum_{i=0}^{K-1} \min \left(FR_i[j] + \lambda \times FA_i[j]\right).$$

In order to find $\min(FR_i[j] + \lambda \times FA_i[j])$, let us suppose that there is a point moving on the ROC curve. The value of $(FR_i[j] + \lambda \times FA_i[j])$ is always changing when the point is moving. Only when the point reaches another line with a slope of λ, $(FR_i[j] + \lambda \times FA_i[j])$ is minimized because all the other points on this ROC curve are to its up-right. So $\min(FR_i[j] + \lambda \times FA_i[j])$ is the same as finding one line, which has a slope of λ and is tangent to the ith ROC curve. By finding each of this line in K ROC curves, we get K pair of FAR and FRR, then average them and we get the corresponding FAR and FRR pair in the combined ROC curve. This algorithm can also be illustrated by Fig. 11. So essentially by tuning different λ, we can get combined ROC curve. Actually, this kind of combining curves method also has applications in many other fields, such as the rate-distortion techniques for video coding [21].

We have carried out two experiments, both of which have the same experimental configuration except using different threshold methods. Here we collect a face data set, which has 13 subjects. Each one has 5 neutral training faces, 70 test images with different expressions. And we use the universal PCA approach in the authentication. From Fig. 14, we can see that better performance can be obtained by combining individual thresholds. This means that in an authentication system with multiple users, instead of choosing a fixed threshold, we should adjust each subject's threshold according to its own ROC curve. By doing this, improvement can be seen in the performance of the whole authentication system.

6. Experiment results

Experiments and evaluations are important parts of any face authentication system. In addition to using publicly available face data sets, we collected different face sets specified for each experiment. Also in order to isolate the effects of facial expression variations and registration errors, we assume that all the training and test images are captured under the same lighting condition.

6.1. The experimental setup

Before discussing the experiment results, we present some details about our algorithm. First, the face region needs to

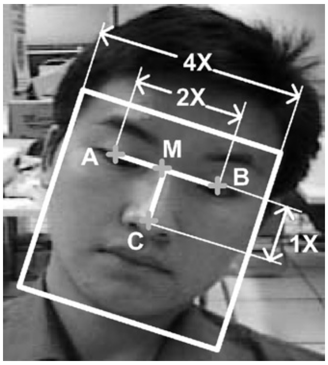

Fig. 12. A face cropping operation.

be cropped from the whole image. It can be automatically accomplished by a cropping operation based on the locations of two eyes, which are obtained from an eye-tracking algorithm [18]. As shown in Fig. 12, the cropping operation consists of the following steps. Firstly, one line connects two eye locations, which is called *eyes line* AB. In the middle of AB, we draw another line called *center line* MC. It is perpendicular to AB and has half of the length. Then the end of the center line, C, is used as the center of the face region to be cropped. The cropped region is a square parallel to the eyes line and the center line, and its size is two times the length of the eyes line.

Given any two training images, we generate the optical flow as shown in Fig. 13. First, the background regions below the cheeks in the face image are removed because the background affects the optical flow calculation, and thus interferes with the authentication. As seen from Fig. 13(b), zero is filled into the two triangle regions in the lower part of the face square. Next, we determine the optical flow using the Lucas–Kanade algorithm [22]. Third, the optical flow is down sampled to be half of its original size in order to speed up the PCA training and to clean up the noisy motion vectors. Finally within this smaller-size optical flow, the background and four side boundaries are removed because usually the boundary does not result in accurate motion estimation in the optical flow algorithm. Now the down-sampled optical flow image can be scanned into a vector, whose dimension is much lower than the unsampled one. For example, if the original image dimension is 64×64, the optical flow vector used for training is only 1620, about 20% of the unsampled one. For the down sampling in two dimensions, there are several methods, such as low pass filtering and median filtering. From the experiment, we found that applying the low pass filter (with the coefficients of $\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}$) to the horizontal and vertical motion vector separately yields good performance.

6.2. Individual PCA versus universal PCA

We collected a training set including 17 subjects, with 51 face images for each subject. Each image has been registered based on the locations of the eyes, and has the same size of 64×64 pixels. By using the leave-one-out method [23], 50 faces of each subject can be utilized for training and one image for testing. Then we replace the test image with one of the training images, and repeat the training and test process. Thus, we can get 51×17 client claims and $51 \times 17 \times 16$ imposter claims.

Based on this experiment configuration, we implemented the universal PCA approach and the individual PCA approach introduced in Section 2. Another task in comparing these two methods is to select the number of eigenvectors to be used. In both implementations, we select the number of eigenvectors such that, the corresponding eigenvalues can capture a fixed percentage of total energy. The two ROC curves are shown in Fig. 15. From it we can see that the individual PCA method provides a better performance.

6.3. Individual eigenflow approach versus PCA approaches

From the previous experiment, we have already seen that the individual PCA works better than the traditional PCA approach for authentication. When we perform experiments involving our eigenflow-based approach, we compare its performance with both individual and universal PCA approaches. In this experiment, we want to show how the new eigenflow-based approach works when both training images and test images have facial expression variations and registration errors. Three cases are concerned: only expression variations, only registration errors, or both.

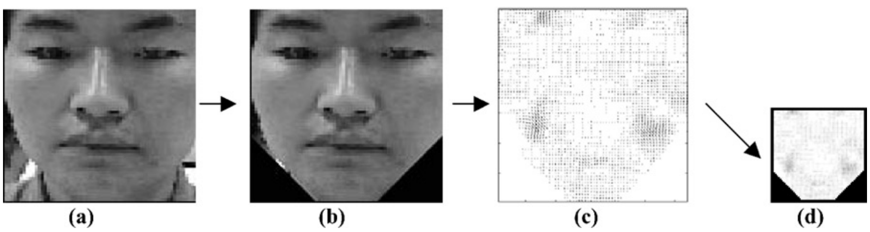

Fig. 13. Four steps of getting a training optical flow: (a) original image, (b) removing the backgrounds, (c) getting optical flow and (d) down-sampling optical flows and removing boundaries.

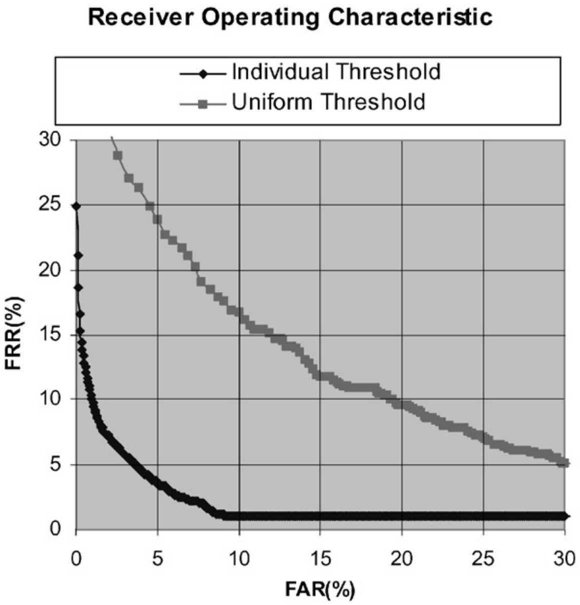

Fig. 14. Comparison of the uniform and individual threshold methods.

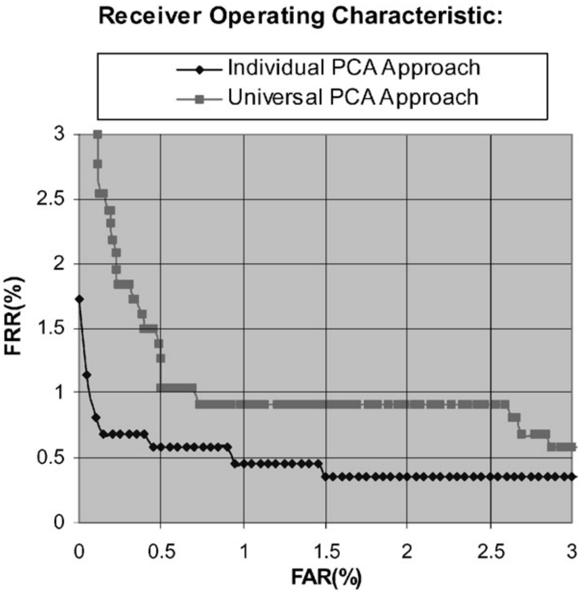

Fig. 15. Performance comparison between the individual PCA and universal PCA approach.

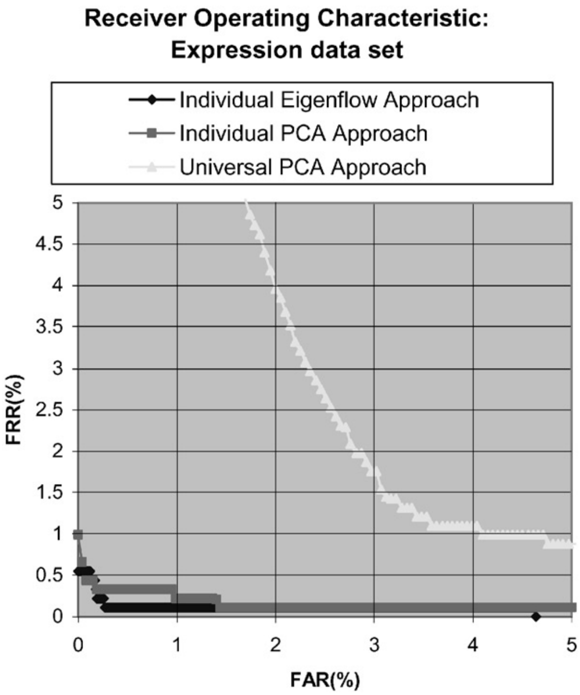

Fig. 16. Experiment results on the data set with expression variations.

6.3.1. Expression variations

The first data set has only expression variations. 13 subjects are included in this set. Each has 5 images for training, and 70 images for testing. The reason we use more test images than training images is that we want to get a smoother ROC curve. Also, only a few images may be available for training in a practical setup. Each of the five training images represents different expressions, such as neutral, happy, angry, sad, and surprise. All of these images are well registered by the locations of the eyes. Here we implement three algorithms: the individual PCA approach, the universal PCA approach, and the individual eigenflow approach. From the result shown in Fig. 16, we can see that for most part of the curve, our approach yields better performance. Also the improvement is significant compared to the universal PCA approach.

6.3.2. Registration errors

The second data set has registration variations for each subject. To synthesize the registration variations, we perturb the eye tracking results to some extent while cropping the face region. Given one well-registered face image, we can synthesize 625 images by cropping face regions based on 25 points around the left eye location, and 25 points around the right eye location. These images are used for training one subject. The same method is used to generate the 625 test images except there is larger offset while selecting the eyes' neighbor points, which means test images have larger registration errors than training images. So all these synthesized training images can represent different kinds of registration errors. As shown in Fig. 17, the eigenflow-based approach has shown much better performance than the PCA approach.

Actually our approach can also work on the data set, which contains real registration errors resulted from imprecise tracking of eye locations in practical applications [18]. We collect 81 training images and 624 test images with real

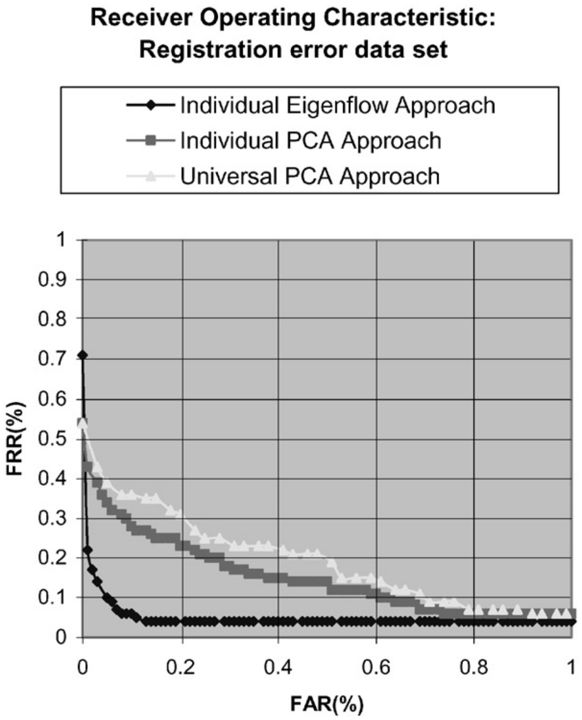

Fig. 17. Experiment results on the data set with registration errors.

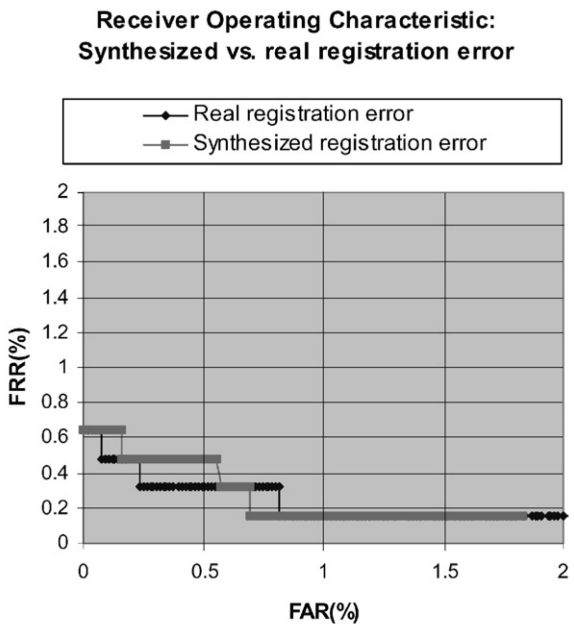

Fig. 18. Authentication for one subject in case of synthesized and real registration errors.

registration errors for one subject. Also we have the same number of images with synthesized registration errors for the same subject. After we perform individual eigenflow approach for this subject, from Fig. 18 we observe that a similar or even better authentication performance can be obtained in the case of real registration errors.

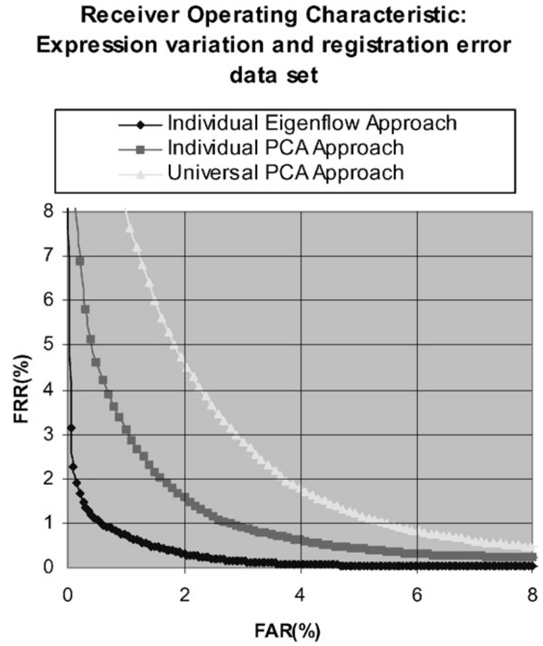

Fig. 19. Experiment results on the data set containing both expression variations and registration errors.

6.3.3. *Expression variations and registration errors*

The third data set has both expression variations and registration errors. First, for each of the 13 subjects, 5 images with different expressions are obtained to be the reference images. Then, for each reference image, 624 images can be synthesized to include all kinds of registration errors. Thus, 3125 images are collected for each subject. We also generate 3120 test images for each subject using the same approach. However, there are two differences. First, all the test images have different expressions compared to the reference images. Second, the test images have larger registration errors than the training images.

Since in this experiment, we have much more test images than in the previous two experiments, the denominator in computing FAR and FRR is much larger. That is why a much smoother ROC curve can be observed from Fig. 19.

6.4. *Common eigenflow approach versus PCA approaches*

As described in Section 4, a lot of data should be collected for training a common eigenflow space. In the meantime, the common weights are also obtained, which indicate how much contribution each residue can provide to the final measurement.

After that, we collected another data set, which has 13 subjects. Each one has 5 neutral training faces, and 70 test

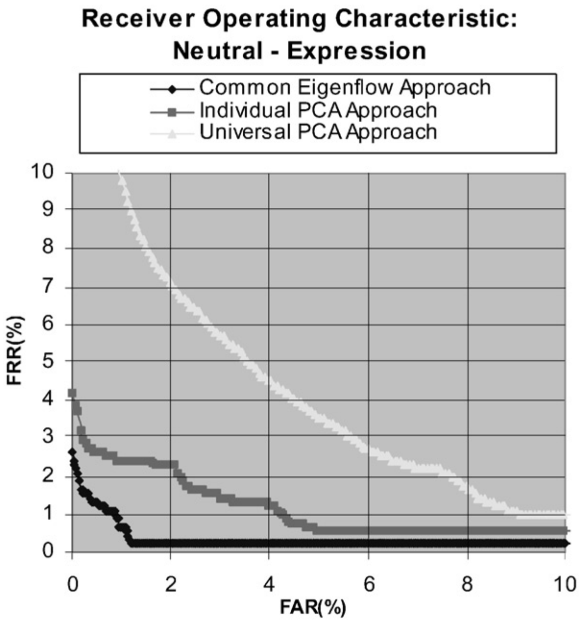

Fig. 20. Experiment results on the neutral training set and the test set with expression variations.

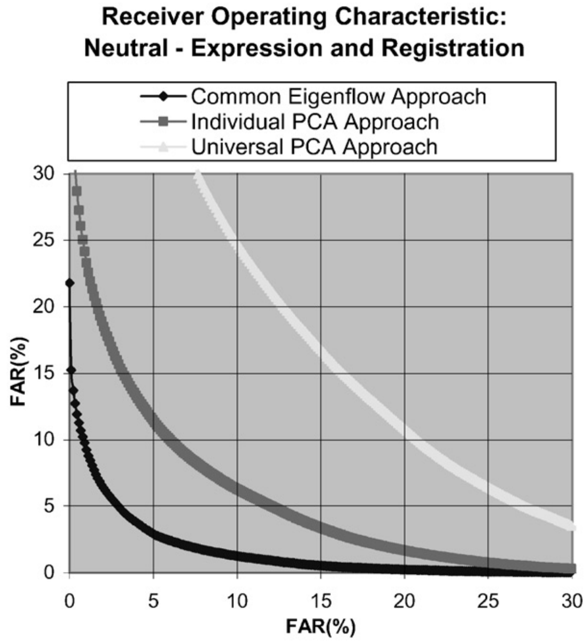

Fig. 21. Experiment results on the neutral training set and the test set with both facial expression variations and registration errors.

images with different expression variations. Then based on the common eigenflow approach, each test image can yield two residue values, which are combined into one measurement by using the common weights. Finally from Fig. 20 we see that the common eigenflow-based approach has better performance than the individual PCA approach. Also it is much better than the universal PCA approach.

In another experiment, the common eigenflow space is trained on a data set, which contains both facial expression variations and registration errors. After that, we collect a test data set with 13 subjects. Each has 3120 test images, which are all different in expressions and registration errors. Using this test set, we evaluate the three approaches based on the same training set with 5 neutral images for every subject, and show the results in Fig. 21. We can see that the common eigenflow approach has much better performance than the other two. Actually, this figure shows the case that will be seen most often in practical applications.

6.5. Face recognition based on the individual PCA

We have shown that the individual PCA outperforms the universal PCA for face authentication. We believe that this it also true for face recognition. In order to show this, we compare results on two public databases: one is the FERET database [24], the other is the ORL database [25], as shown in Fig. 22. As stated before, authentication requires only one process for one specific subject, whereas for recognition, the same process is repeated for every subject. Thus, every authentication approach can also be used for recognition.

We extract a subset, which has 21 subjects, from the FERET database. Each subject has 10 images. For both the individual PCA and the universal PCA approach, we use the first 5 images for training, and the remaining 5 images for testing. We found the error rate of the individual PCA is about 12%, but the error rate of the universal PCA is 20%. In the ORL database, there are 40 subjects, each with 10 images. We use the same test scheme as the FERET data set. The error rate of the individual PCA is about 5%. However, the error rate of the universal PCA is 12%. From these experiments, we can see that the individual PCA can achieve significant improvement in face recognition compared to the universal PCA.

7. Conclusions and discussion

In this paper, we introduced the eigenflow as a novel approach for face authentication tolerant to expression variations and registration errors. Being an intensity-based approach, the eigenface method is very sensitive to expression variations and registration errors. To solve this problem, we proposed to use optical flow to determine the visual motion between pairs of face images. Then PCA is applied to these optical flows to produce eigenvectors, called eigenflows. These eigenflows can represent the variations appearing in training images. Based on the individual eigenflow space, two kinds of residues, the eigenflow residue and the optical flow residue, are combined into the authentication measurement with LDA. We showed that, in

Fig. 22. The set of 10 images from one subject in the ORL database. Considerable variations can be seen in this set.

Table 1
Application cases and our approaches

Training data	Test data	Approach (Section)
Neutral	Neutral	Individual PCA (2.2)
Neutral	Registration errors	Individual eigenflow (3)
Neutral	Expression variations	Common eigenflow (4)
Neutral	Expression variations and registration errors	Common eigenflow (4)
Expression	Expression variations	Individual eigenflow (3)
Expression	Expression variations and registration errors	Individual eigenflow (3)

The number following the approach name in the third column is the section number in this paper.

general, individual PCA is more suitable for classification compared to universal PCA. We also proposed an approach to optimizing and evaluating the system performance of an individual space-based approach for multiple subjects. Different face data sets are collected for experiments. As shown in Table 1, in all these cases, we have seen that the eigenflow-based approach exhibits better performance than the traditional PCA method.

Although we applied our algorithm to face authentication, it can be used for face recognition as well. Essentially, face recognition can be accomplished by applying authentication to the unknown image with respect to each subject in the data set and then choosing the one with the smallest difference. Also both the performance of authentication and the performance of recognition rely on the discrimination of personal identity, which is well modeled in our approach.

From the experiments, we found that the optical flow algorithm we used did not work very well in the case of abrupt motion or large registration errors, which could also be seen from the residue image. Even with this difficulty, the eigenflow space still models the dominant variations and shows good discrimination in the authentication performance. In the future, we will seek better optical flow algorithms. With a better optical flow algorithm, both residues, the eigenflow residue and the optical flow residue, can be improved in the sense of discriminative ability.

The basic idea proposed in our approach is that by explicitly modeling the "difference" between face images under different variations, better classification performance could be achieved. Our method can also be extended to model other variations that appear in faces, such as illuminants and poses.

Acknowledgements

This work is supported in part by Pennsylvania Infrastructure Technology Alliance (PITA) and Institute for Information Industry (III). The authors would like to thank Dr. Yingli Tian for helping evaluating the performance of different optical flow methods. Thanks to Dr. Simon Baker, Dr. Henry Schneiderman, and Dr. Susan M. Thornton for fruitful discussion. Thanks to a number of volunteers in Electrical and Computer Engineering Department of Carnegie Mellon University who helped us collect the face databases specific for this research.

References

[1] T. Kanade, Picture processing system by computer complex and recognition of human faces, Doctoral dissertation, Department of Information Science, Kyoto University, November 1973.

[2] R. Chellappa, C.L. Wilson, S. Sirohey, Human and machine recognition of faces: a survey, Proc. IEEE 83 (5) (1995) 705–741.

[3] C.L. Kotropoulos, A. Tefas, I. Pitas, Frontal face authentication using discriminating grids with morphological feature vectors, IEEE Trans. Multimedia 2 (1) (2000) 14–26.

[4] P.N. Belhumeur, J.P. Hespanha, D.J. Kriegman, Eigenfaces vs. Fisherfaces: recognition using class specific linear projection, IEEE Trans. Pattern Anal. Mach. Intell. 19 (7) (1997) 711–720.

[5] R. Brunelli, T. Poggio, Face recognition: features versus templates, IEEE Trans. Pattern Anal. Mach. Intell. 15 (10) (1993) 1042–1052.

[6] T. Fromherz, P. Stucki, M. Bichsel, A survey of face recognition, MML Technical Report No. 97.01, Department of Computer Science, University of Zurich, Zurich, 1997.
[7] M. Turk, A. Pentland, Eigenfaces for recognition, J. Cognitive Neurosci. 3 (1) (1991) 71–86.
[8] A. Pentland, B. Moghaddam, T. Starner, View-based and modular eigenspaces for face recognition, Technical Report 245, MIT Media Lab Vismod, 1993.
[9] J.J. Lien, A. Zlochower, J.F. Cohn, T. Kanade, Automated facial expression recognition, Proceedings of the Third IEEE International Conference on Automatic Face and Gesture Recognition, Nara, Japan, April, 1998, pp. 390–395.
[10] Y. Yacoob, L.S. Davis, Recognizing human facial expressions from long image sequences using optical flow, IEEE Trans. Pattern Anal. Mach. Intell. 18 (6) (1996) 632–646.
[11] P. Kruizinga, N. Petkov, Optical flow applied to person identification, Proceeding of the 1994 EUROSIM Conference on Massively Parallel Processing Applications and Development, Delft, The Netherlands, Elsevier, Amsterdam, 21–23 June, 1994, pp. 871–878.
[12] B. Moghaddam, T. Jebara, A. Pentland, Bayesian face recognition, Pattern Recognition 33 (2000) 1771–1782.
[13] R.O. Duda, P.E. Hart, D.G. Stork, Pattern Classification, 2nd Edition, Wiley, New York, 2001.
[14] R.A. Horn, C.R. Johnson, Matrix Analysis, Cambridge University Press, Cambridge, 1985.
[15] H. Murakami, B.V.K.V. Kumar, Efficient calculation of primary images from a set of images, IEEE Trans. Pattern Anal. Mach. Intell. 4 (5) (1982) 511–515.
[16] E. Oja, Subspace Methods of Pattern Recognition, Wiley, Letchworth, Hertfordshire, England, New York, 1983.
[17] C.L. Fennema, W.B. Thompson, Velocity determination in scenes containing several moving objects, Comput. Graphics Image Process. 9 (1979) 301–315.
[18] F.J. Huang, T. Chen, Tracking of multiple faces for human–computer interfaces and virtual environments, IEEE International Conference on Multimedia and Exposition, New York, July 2000.
[19] K. Fukunaga, Statistical Pattern Recognition, Academic Press, New York, 1989.
[20] W. Zhao, R. Chellappa, P.J. Phillips, Subspace linear discriminant analysis for face recognition, Technique Report, CS-TR-4009, University of Maryland, College Park, April 1999.
[21] A. Ortego, K. Ramchandran, Rate-distortion methods for image and video compression, IEEE Signal Process. Mag. 15 (6) (1998) 23–50.
[22] B.D. Lucas, T. Kanade, An iterative image registration technique with an application to stereo vision, Proceedings of the DARPA IU Workshop, April, 1981, 121–130.
[23] K. Fukunaga, D. Kessell, Estimation of classification error, IEEE Trans. Comput. C-20 (1971) 1521–1527.
[24] P.J. Phillips, M. Hyeonjoon, S.A. Rizvi, P.J. Rauss, The FERET evaluation methodology for face-recognition algorithms, IEEE Trans. Pattern Anal. Mach. Intell. 22 (10) (2000) 1090–1104.
[25] AT& T Laboratories Cambridge, The ORL Database of Faces, http://www.cam-orl.co.uk/facedatabase.html.

About the Author—XIAOMING LIU received the B.E. degree from Beijing Information Technology Institute, Beijing, China, and the M.E. degree from Zhejiang University, Hangzhou, China, in 1997 and 2000, respectively, both in Computer Science. Since January 2000, he has been working towards his Ph.D. degree in Electrical and Computer Engineering at Carnegie Mellon University, Pittsburgh, PA. His research interests are in the areas of image processing, pattern recognition and computer vision. He is a student member of the IEEE.

About the Author—TSUHAN CHEN received the Ph.D. degree in Electrical Engineering from the California Institute of Technology, Pasadena, California, in 1993. From August 1993 to October 1997, he worked in the Visual Communications Research Department, AT& T Bell Laboratories, Holmdel, New Jersey, and later at AT& T Labs-Research, Red Bank, New Jersey, as a senior technical staff member and then a principle technical staff member. Since October 1997, he has been with the Electrical and Computer Engineering Department, Carnegie Mellon University, as an Associate Professor. His research interests include multimedia signal processing and communication, audio-visual interaction, video coding and multimedia standards.

Dr. Chen co-founded and chaired the Multimedia Signal Processing Technical Committee of the IEEE Signal Processing Society. He serves as Associate Editor for *IEEE Trans. on Image Processing* and *IEEE Trans. on Multimedia*. He serves on the Steering Committee of *IEEE Trans. on Multimedia* and the Editorial Board of *IEEE Signal Processing Magazine*. He was a technical co-chair of the First IEEE Conference on Multimedia and Expo in 2000. He has recently been appointed as the Editor-in-Chief for *IEEE Transactions on Multimedia* for 2002–2004. Dr. Chen has published tens of technical papers and holds eleven US patents. Dr. Chen is a recipient of the National Science Foundation CAREER Award.

About the Author—B.V.K. VIJAYA KUMAR received his B.Tech., and M.Tech., degrees in Electrical Engineering from Indian Institute of Technology, Kanpur and his Ph.D. in Electrical Engineering from Carnegie Mellon University (CMU), Pittsburgh. Since 1982, he has been a faculty member in the Department of Electrical and Computer Engineering at CMU where he is now a Professor. Professor Kumar also served as the Associate department head from 1994 to 1996. Professor Kumar's research interests include Pattern Recognition, Image Processing, and Coding and Signal Processing for Data Storage Systems. He has authored or co-authored seven book chapters and more than 260 technical papers in these areas in various conference proceedings and journals. He served as a Topical Editor for the Information Processing division of *Applied Optics.* He guest-edited the September 1990 special issue of *Optical Engineering* devoted to *Optical Pattern Recognition* and the April 2001 issue of *IEEE Journal of Selected Areas of Communication* (JSAC) devoted to *Signal Processing for High-Density Storage Channels.* Professor Kumar has served on many conference program committees including the SPIE conference on *Optical Pattern Recognition*, the *Optical Data Storage* (ODS) conference, *The Magnetic Recording Conference* (TMRC) and the *International Symposium on Optical Memories* (ISOM). Professor Kumar is a senior member of IEEE, a Fellow of SPIE—The International Society of Optical Engineering, and a Fellow of Optical Society of America (OSA). He is listed in various professional Who's Whos including Marquis' *Who's Who in Science and Engineering* and *Who's Who in the World* and in the *American Men and Women of Sciences.*