

We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

4,800

Open access books available

122,000

International authors and editors

135M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com



A New Multimodal Biometric for Personal Identification

Pohsiang Tsai, Tich Phuoc Tran and Longbing Cao

*Faculty of Engineering and Information Technology – University of Technology, Sydney
Australia*

1. Introduction

Biometric technology using a single human body characteristic such as face, gait or voice has gained immense attention with successful applications in video surveillance. Furthermore, facial recognition has been recognized as the least intrusive technology that can be implemented in many places without hazardous problems. Though existing biometric systems have been reported to be effective under certain conditions, there is a great need to improve their recognition performance. Possible alternatives to current approaches include the use of different biometric information or the combination of different biometric sources. Relevant literature indicates the possibility of using facial behavior as another behavior biometric cue. As this biometric cue reflects the internal dynamic changing factors of an individual, it also plays an important role just as the face biometric does in video surveillance. Moreover, existing research in appearance-based facial recognition always addresses the common problem of within-class variations under illumination and poses and/or facial expressions which degrades the recognition performance. Most of the algorithms (Belhumeur et al., 1997) (Chen et al., 2000) (Lu et al., 2003c) (Lu et al., 2003b) (Lu et al., 2003a) (Juwei et al., 2003) (Lu et al., 2005) (Kong et al., 2005) in facial recognition are developed to cope with the singularity problem in the presence of these variations. Some papers (Martinez, 2000) (Liu et al., 2002) (Liu et al., 2003) (Bronstein et al., 2003) (Bronstein et al., 2007) even consider facial expressions as noise that will degrade the system performance, and they attempt to build robust systems that are invariant to these variations. However, no research has been conducted to see if these intra-personal variations, especially under facial expression changes or dynamic changes of intra-personal information could help the extra-personal separation. Can within-class variation help between-class separation? Or can intra-personal facial expression variations assist extra-personal separation? We assume that “intra-personal facial expression variations could assist extra-personal separation.”

In the following section, we give an overview of multimodal biometrics and soft biometrics. We discuss how multiple biometrics and soft biometric can improve classification performance. We then discuss related work that uses other biometrics in combination with facial biometric at a distance and our proposed fusion framework. Finally, we give the experimental results and conclusion.

2. Overview of Multimodal Biometrics

Biometric systems have their success in single modalities; however, there are still opportunities to improve their limitations. Jain et al. (Jain et al., 2004b) stated that these limitations include the noise in the sensed data, intra-personal variations, distinctiveness of inter-personal variations, non-universality, and spoof attacks. Furthermore, Faundez-Zanuy (Faundez-Zanuy, 2005) stated the main drawbacks in a different uni-modal biometric system. For example, in fingerprint biometrics, some people might not have their fingerprint characteristics because they were old, their fingerprints might be too oily, dry, wet or warm; or their fingerprints might be damaged temporally or permanently. These situations make it difficult or impossible for scanners to acquire their fingerprints. Facial biometrics' weaknesses are due to variations in makeup, pose or illumination variations. Therefore, it is necessary to explore other sources for performance improvement so as to achieve robustness under various circumstances.

An interesting question is "Can multiple biometrics improve performance?" Hong et al. (Hong et al., 1999) examined the possible performance improvement of a biometric system if this system integrates multiple biometrics. They proved that by integrating with other multiple biometric sources, the performance was indeed improved. Multiple biometrics also makes spoofing more difficult because of the difficulty of simultaneously spoofing multiple biometric characteristics (Jain et al., 2006).

These possible schemes of combining various biometric cues in multimodal biometrics include the fusion of 2D and 3D face images (Chang et al., 2005); the fusion of 2D image, lip motion, and voice (Fröba et al., 2000); and the fusion of 2D image, 3D facial shape and infrared facial heat pattern image (Chang et al., 2004), speech and face (Brunelli and Falavigna, 1995), faces and fingerprints (Lin and Anil, 1998), face and gait (Shakhnarovich et al., 2001). It is clear that a biometric system combining several features can achieve better recognition accuracy than a single-feature system.

Hu et al. (Hu et al., 2004) explained the three reasons for using multimodal biometrics. Firstly, it can be used in some highly secured places such as military bases or government units for visual surveillance purposes. A security system based on this can store human biometric cues such as facial appearance, gait or height in the database for access control purposes. Secondly, a police station can setup a visual surveillance system that stores criminals' biometric features in railway stations or casinos and monitor the criminal activity at a distance. Moreover, combining the various biometric cues such as facial appearance and gait can provide a more reliable alarm to the police if the facial appearance quality is low. Thirdly, it can provide reliable object recognition of either human or inanimate objects.

3. Soft Biometrics

Any physiological and behavioral human biometric cues can be used for personal identification, and these characteristics are also called 'hard biometrics'. Other human secondary information, such as gender, ethnicity, age, height, and weight and eye color are called 'soft biometrics'. These have the potential of complimenting the personal identification information provided by any hard biometrics (primary sources) and improving recognition performance (Jain et al., 2004a); however, the ancillary information may not be distinctive and permanent enough to differentiate any two individuals. "Can soft biometric traits assist user recognition?" Jain and Dass (Jain et al., 2004a) studied their

usability with an example of using a fingerprint biometric system for this purpose. They provided additional personal information (gender, ethnicity, and height) to the fingerprint biometric systems and showed that it increased recognition rates about 6%.

Fig. 1 shows some examples of soft biometrics and Fig. 2 shows the general framework for integrating soft biometrics with hard biometrics.

Soft biometrics cannot be solely used for reliable personal identification because they are not distinctive and permanent. They can complement the primary biometric system. They also can be extended to multiple biometric systems. For example, Jain et al. (Jain et al., 2004a) proposed a hybrid biometric system that uses face and fingerprint as the primary biometrics and integrated with secondary biometrics (e.g. gender, ethnicity, and height). They found that the complementary secondary biometrics can help improve recognition performance significantly.



Fig. 1. Some soft biometric examples (Jain et al., 2004a)

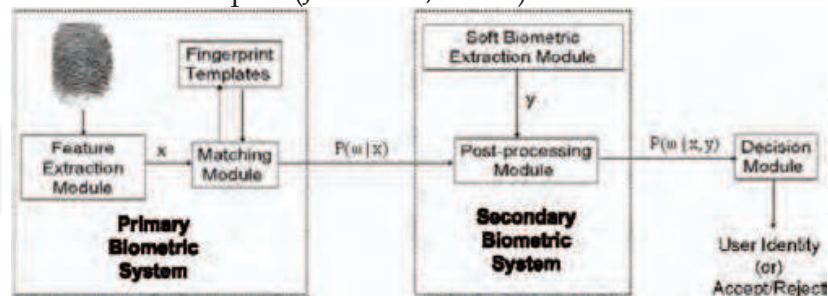


Fig. 2. General framework for soft biometric integration with hard biometrics (Jain et al., 2004a).

In the following section, we will discuss each level of fusion techniques in more detail.

4. Level of Biometric Fusion

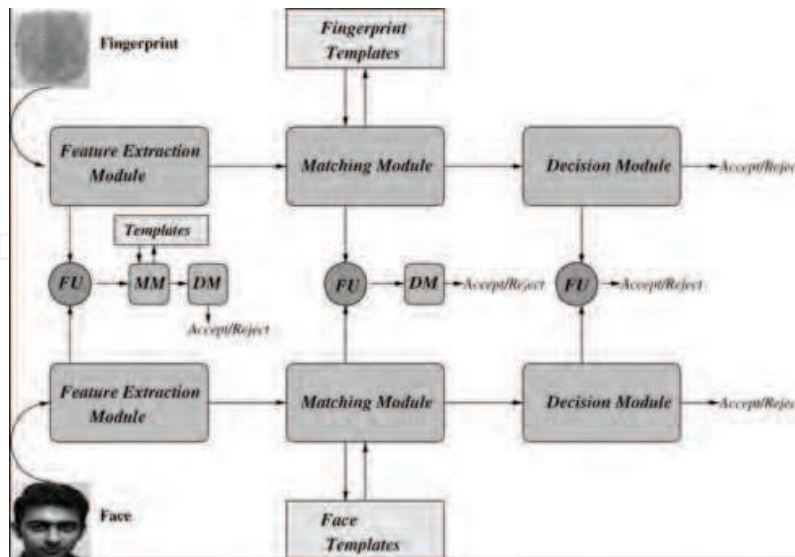


Fig. 3. The three different fusion levels in a bimodal biometric system (Ross and Jain, 2003).

The three levels of fusion techniques are described as follows:

5. Fusion at the feature extraction level

Individual biometric cues extracted from their own feature extraction module are concatenated into a larger feature vector in the fusion module. A further treatment of using feature selection or extraction methods on this new combined vector is necessary so as to avoid the curse-of-dimensionality problem.

Fusion at the feature level is more effective than fusion in other levels because more salient features are obtained. However, it has its drawbacks which include a highly correlated relationship between features which need to be removed, the high dimensional problem of the combined vectors, and privacy issues which hinder the availability of collecting multimodal biometric data (Jain et al., 2006).

6. Fusion at the matching score level

Individual matchers provide their similarity scores to indicate the proximity of the input with the template in the trained database. The scores are normalized using transformation techniques to transform the scores of the individual modalities into a common domain and an appropriate fusion strategy to combine the transformed scores (Jain et al., 2006). This level is also called confidence or rank level. These techniques include Min-max, Z-score, or Tanh.

7. Fusion at the decision level

Individual matchers provide their own decisions based on their own feature vectors in the feature extraction module. The final decision is then made based on each individual decision by using techniques such as majority voting, behavior knowledge space, weighted voting based on the Dempster-Shafer theory of evidence, and AND and OR rules. This level is also called the abstract level.

7.1 Related Work

Uni-modal facial biometric is an un-intrusive collection of human facial characteristic that is also an important component of multiple biometric systems. This biometric can integrate with other (many) biometrics to provide a more robust identification system. A special issue in 2007 presents recent advances in biometric systems (Boyer et al., 2007) that addresses the advances in uni-modal and multimodal biometrics. It concludes that most human biometric cues such as face, fingerprint, voice, signature, and iris are large active areas while other biometric cues such as gait, ear, brain signal recordings and infrared imaging of hand vein patterns are the newer and smaller areas in uni-modal biometrics. In multiple biometrics, face and gait fusion and 2D+3D face fusion are the two active research areas.

We present literature that uses other biometrics in combination with facial biometrics at a distance. Chang et al. (Chang et al., 2003) combined the face and ear for multiple biometrics. They used the PCA method applied to these two biometrics (called Eigen-Faces and Eigen-Ears). Their results showed the potential performance increase in combination with face and ear that out-perform either face or ear. However, their limitation is that ear images were only in profile view. Chang et al. (Chang et al., 2005) fused 2D and 3D facial images for multimodal biometrics. They used the PCA method for salient feature extraction of individual modalities and score normalization method for normalizing multiple matcher scores. Their results showed that the fused 2D and 3D facial images performed better than using either 2D or 3D facial images alone. Shakhnarovich et al. (Shakhnarovich et al., 2001) integrated face and gait from multiple views for multimodal biometrics. They used a view-normalization approach to multiple views. Then, each biometric trait's matching score was combined. Their results showed the integrated face and gait biometrics had improved the recognition performance over either one alone. Zhou and Bhanu (Zhou and Bhanu, 2007) presents the latest trend in face and gait fusion. They tackled the challenging problem of using shape and intensity information of side view of the face combined with side gait information. Their approach was rather different to the traditional approaches of using a frontal view for face and a side view for gait. They used Enhanced Side Face Image (ESFI) approach and Gait Energy Image (GEI) approaches for integrating face and gait for non-cooperating human recognition at a distance. Liu and Sarkar (Liu and Sarkar, 2007) explored the possibility of using both face and gait in an outdoor environment at a distance. Toh (Toh et al., 2008) et al. proposed a weighted power series model to minimize an approximated error rate formulation for visual and infra-red fusion for facial verification. Infrared facial imaging is an additional option for multiple biometrics combining with face biometric. However, as this biometric cue requires a costly sensor, it is not applicable to a low-cost camera technology in vision-based surveillance at a distance. All the research concluded that "the performance of recognition is improved when combining multiple biometrics. Multiple biometric combination outperformed either single biometric."

In the following section, we will propose a fusion framework that integrates facial appearance and facial expression features.

8. Proposed Framework

We introduce a face coding and recognition method, the Fisher's Linear Discriminant Classifier (FLDC), which employs the FLD Model (FLDM) (Duda et al., 2000a) for integrated facial appearance and facial expression features. The facial appearance provides textural information of a face, while the facial expression changes are geometrically encoded by distance-based facial fiducial points that capture the changes of a face when it displays facial expressions or emotions. To reduce the dimensionality of the original facial appearance and facial expression spaces, PCA, constrained by the FLDM for enhanced discriminant capability, derives low dimensional features, which are then combined using a normalization procedure in order to form integrated features accounting for both facial appearance and facial expression information. Finally, the integrated features are processed by the FLDM for face recognition.

Fig. 4 shows the proposed framework for PCA-level feature fusion. The gray level intensity of facial appearance images and fiducial point distance-based measurement of facial expression images are extracted firstly by PCA. Note that we only show the neutral expression of facial appearance and facial expression images in Fig. 4.

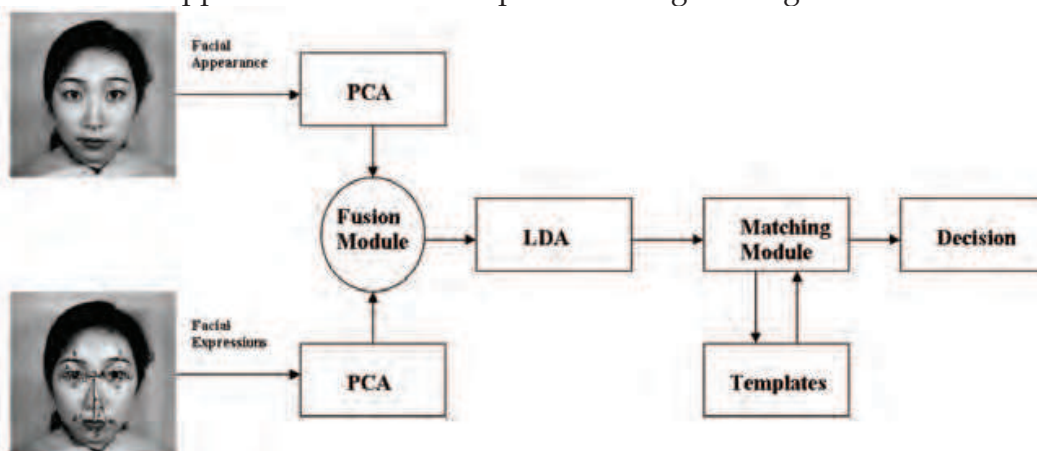


Fig. 4. The framework for PCA-level feature fusion.

For other expression images, they go through the same procedures as shown in Fig. 4. After PCA extraction, the extracted two different biometrics are then combined together to form larger fused vectors. These fused vectors are then applied to LDA to get reduced dimensionality and the increased discriminant data distribution. Finally, the feature obtained by LDA is compared with the matching template before making the decision. In the following sections, we discuss the feature extraction techniques used in the proposed framework.

8.1 Principal Component Analysis

Let a facial appearance or expression image X_i be a vector of dimension d . Denote the training set of n facial appearance or expression images by $X = (X_1, X_2, \dots, X_n) \subset \mathbb{R}^{d \times n}$,

and we assume that each image belongs to one of C classes. Define the covariance matrix X as follows :

$$\Sigma_X = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})(X_i - \bar{X})^T \quad (1)$$

where $\Sigma_X \in \mathbb{R}^{d \times d}$ and $\bar{X} = (1/n) \sum_{i=1}^n X_i$. Then, the eigenvalues and eigenvectors of the covariance matrix Σ_X are calculated. Let $\Phi = (\phi_1, \phi_2, \dots, \phi_r) \in \mathbb{R}^{d \times d}$ ($r \leq d$) be the r eigenvectors corresponding to the r largest eigen-values. Thus, for a set of original facial appearance or expression images $X \in \mathbb{R}^{d \times n}$, their corresponding eigen-based feature $Y \in \mathbb{R}^{r \times n}$ can be obtained by projecting X onto the eigen-based feature space Φ as follows:

$$Y = \Phi^T X \quad (2)$$

Hence, this reduced lower dimensional vector Y captures the most expressive features of the original data X .

8.2 Fisher Linear Discriminant Analysis

Let a facial appearance or expression image X_i be a vector of dimension d . Denote the training set of n facial appearance or expression images by $X = (X_1, X_2, \dots, X_n) \in \mathbb{R}^{d \times n}$, and we assume that each image belongs to one of C classes. Define the between-class scatter and the within-class scatter matrices as follows {Bishop, 1995 #91}:

$$S_B = \sum_{i=1}^C n^i (\bar{X}^i - \bar{X})(\bar{X}^i - \bar{X})^T \quad (3)$$

$$S_W = \sum_{i=1}^C \sum_{X_k \in X_i} (X_k - \bar{X}^i)(X_k - \bar{X}^i)^T \quad (4)$$

where $\bar{X} = (1/n) \sum_{j=1}^n X_j$ is the mean image of input vectors, and $\bar{X}^i = (1/n^i) \sum_{j=1}^{n^i} X_j^i$ is the mean image of the i th class, n^i is the number of samples in the i th class and C is the number of classes. Therefore, if S_W is nonsingular, the optimal projection W_{opt} is chosen as the matrix with orthonormal columns which maximizes the ratio of the determinant of the between-class scatter matrix of the projected input samples to the determinant of the within-class scatter matrix of the projected input samples. The optimal projection W_{opt} is defined as follows (Kirby and Sirovich, 1990) (Duda et al., 2000b):

$$W_{opt} = \arg \max_W \left| \frac{W^T S_B W}{W^T S_W W} \right| = [W_1, W_2, \dots, W_m] \quad (5)$$

where $\{W_i \mid i = 1, 2, \dots, c-1\}$ is the set of generalized fisher-vectors of S_B and S_W corresponding to the $c-1$ largest generalized fisher-values $\{\lambda_i \mid i = 1, 2, \dots, c-1\}$, i.e.,

$$S_W^{-1} S_B W = \lambda_i W \quad (6)$$

Thus, for a set of original facial appearance or expression images $X \in \mathbb{R}^{d \times n}$, their corresponding fisher-based feature $Q \in \mathbb{R}^{m \times n}$ can be obtained by projecting X onto the fisher-based feature space W_{opt} as follows:

$$Q = W_{opt}^T \cdot X \quad (7)$$

Hence, this reduced lower dimensional vector Q captures the most discriminant features of the original data X .

8.3 Fusion of Facial Appearance and Facial Expression Features

Let $X_1 \in \mathbb{R}^{d_1 \times n_1}$ and $X_2 \in \mathbb{R}^{d_2 \times n_2}$ represent the facial appearance and expression features, where d_1 and d_2 are the number of each representation's dimensionality, and n_1 and n_2 are the number of each representation, respectively. Using (1) one can derive the two covariance matrices, $\Sigma_{X_1} \in \mathbb{R}^{d_1 \times d_1}$ and $\Sigma_{X_2} \in \mathbb{R}^{d_2 \times d_2}$, and the reduced two eigenvector matrices, $\Phi_1 \in \mathbb{R}^{d_1 \times r_1}$ and $\Phi_2 \in \mathbb{R}^{d_2 \times r_2}$, of the facial appearance and expression features, respectively. Finally, one can use (2) to derive the reduced lower dimensional subspaces $Y_1 \in \mathbb{R}^{r_1 \times n_1}$ and $Y_2 \in \mathbb{R}^{r_2 \times n_2}$, of the facial appearance and expression features, respectively, to improve the generalization performance of the FLDC classifier.

The reduced lower dimensional subspaces of Y_1 and Y_2 are then integrated by using (8) of its normalization procedure to form the encoded information of these two facial appearance and expression features (Wechsler, 2007). Note that each reduced subspace is normalized to

have unit norms before they are concatenated to form an augmented combined feature vector. Hence, the new fused feature matrix is defined as follows¹:

$$U = \left(\frac{Y_1^t}{\|Y_1\|} \frac{Y_2^t}{\|Y_2\|} \right)^t \in \mathbb{R}^{r_1+r_2} \quad (8)$$

This new fused feature matrix U then replaces X in Section 0 to yield the most expressive and discriminant lower dimensional subspace Q as in (7).

8.4 Facial Recognition

The FLDC method employs the FLDM on the integrated facial appearance and expression features. When an unknown face image is presented to the FLDC classifier, the original facial appearance and expression features are firstly projected to the Y_1 and Y_2 subspaces as explained in Section 0. Then, the integrated feature U_{test} of the unknown facial image is derived using (8). Let W_{opt} be the final optimal discriminant basis matrix of the FLDA as defined by (7). The new feature space P of the unknown face image is derived as follows:

$$P = W_{opt}^t U_{test} \quad (9)$$

Let Q_c^i be the i th training data point of the class c after the FLDM transformation. The nearest neighbor classification rule is then specified as follows:

$$\|P - Q_c^i\| = \min_{j,k} \|P - Q_k^j\|, \quad P \in \omega_c \quad (10)$$

The unknown facial image is finally classified to class ω_c to whom the unknown facial feature P is the nearest neighbor.

¹ This method assumes the facial appearance and expression features have equal important discriminant information. According to the discussion in Chapter 5 that any human body part can be used as a biometric trait as long as it satisfies the requirement that any individual exhibits distinctive characteristics. Hence, we assume that facial expression features have equal important discriminant information as successful facial appearance-based face recognitions do.

9. Experimental Results

We evaluate the performance of the fusion framework for bi-modal biometrics using facial appearance and facial expression features. The classification performance of the framework is then compared with the uni-modalities; namely, EigenFaces, FisherFace, and the distance-based facial expression features using PCA and LDA methods; we called them EigenExpression and FisherExpression, respectively.

9.1 Experimental Design

The original images in the JAFFE database are raw facial images that include not only the facial information (neutral, angry, disgust, fear, happy, sad, and surprise), but also irrelevant information for facial recognition system. This relevant information includes hair, neck, shoulder, clothes, and background, as shown in Fig. 6.



Fig. 5. Facial Expression Images from JAFFE database [74].

We use the JAFFE database for both appearance-based facial features and distance-based facial expression features. For the facial appearance features, we follow the recommended preprocessing steps (Phillips et al., 2000) so as to avoid incorrect evaluations. These steps are 1) images are registered so that the centers of the eyes are placed on specific pixels; 2) images are cropped and resized (to size 16 x 13) so as to remove the relevant non-face portions and to reduce the high dimensionality; 3) histogram equalization is performed in the resized and cropped facial pixels; 4) the final preprocessed facial data is further normalized to have zero mean and unit variance. Fig. 6 depict examples after the preprocessing stage. Finally, each image is represented as a column vector of $d = 208$.



Fig. 6. Examples after preprocessing.

We design four different experiments for testing our hypothesis as follows:

- I. Neutral images in the training set and facial expression images in the testing set.

Training set: Only neutral images (30 images)

Testing set: All facial expression images (183 images)

- II. Facial expression images in the training set and neutral images in testing set.

Training set: All expression images (183 images)

Testing set: Only neutral images (30 images)

- III. Using 10-fold and leave-one out cross validation for single modality. Neutral and facial expressions both are in the training and testing set.
- IV. Using 10-fold and leave-one out cross validations for bi-modalities. Neutral and facial expressions both are in the training and testing set.

Particularly, in experiment III and IV, following standard performance evaluation practices (Duda et al., 2000a), the facial expression database H , contains N images, is randomly partitioned into two subsets by using a K -fold cross-validation method: The training set Z and the testing set T . That is, the original database H is divided into K partitions². In every run³, the i th partition is used for training set Z and the remaining partitions are combined together to form the testing set T (e.g. $T = H - Z$). Any FR method evaluated here is first trained with Z , and the nearest neighbor classifier is then applied to T , so as to produce the recognition rate, which is defined as $1 - CER$ (classification error rate). Finally, all the recognition rates reported below are averaged over K runs to enhance the statistical accuracy of the assessment. Note that the nearest neighbor classifier is used for the four experiments.

² $K = 10$ (10-fold cross-validation); $K = N$ (leave-one-out cross-validation).

³ There are 10 runs in total because of 10 folds.

9.2 The Facial Image Analysis

9.2.1 EigenFaces Versus FisherFaces

Fig. 7 and Fig. 8 shows the first 28 basis vectors of EigenFaces and FisherFaces of the facial appearance images, respectively. Each eigen-feature based faces are extracted in order to find the optimal either informative or discriminant projection axes, in which data for images lie in linear subspace. One can see from these two figures that while EigenFaces aim to extract the richest information, FisherFaces aim to extract the most discriminant classification. Note that the last five basis vectors of FisherFaces in Fig. 8 look similar and look quite different among others. One possible cause could be the training eigen-values that encode noise.



Fig. 7. First 28 basis vectors of EigenFaces of the facial appearance images.



Fig. 8. First 28 basis vectors of FisherFaces of the facial appearance images.

9.2.2 Linear Subspace Analysis for Facial Appearance and Facial Expression Features

The input facial appearance and facial expression features were normalized to have zero mean and unit variance before applying them to the linear subspace analysis methods. Class labels in the following figures stand for individuals. The normalized input data were then used for extracting the first 30 and 17 dimensions of facial appearance and expression features. Note that by convention we only display the first two components of the transformed matrix of 10 individuals.

Fig. 9 demonstrates that the resultant data distribution of Eigenfaces was overlapped, yet still shows some insight on providing distinctively expressive information in the presence of facial expressions. The resultant analysis is also reflected by Eigenfaces as shown in Fig. 7. Clearly, the finding indicates that even though the resultant data distribution of the Eigenfaces in the presence of facial expression is somewhat overlapped, Fig. 7 and Fig. 9 still tend to reveal that the Eigenface method while providing the expressive information does preserve some discriminant information of individuals.

Fig. 10 demonstrates that the resultant data distribution of Fisherfaces was more overlapped than that of Eigenfaces (Fig. 9). The finding indicates that this method actually smears the classes together so that they are no longer linearly separable in the projected space. Also, this resulting projection space is not simplified for classification.

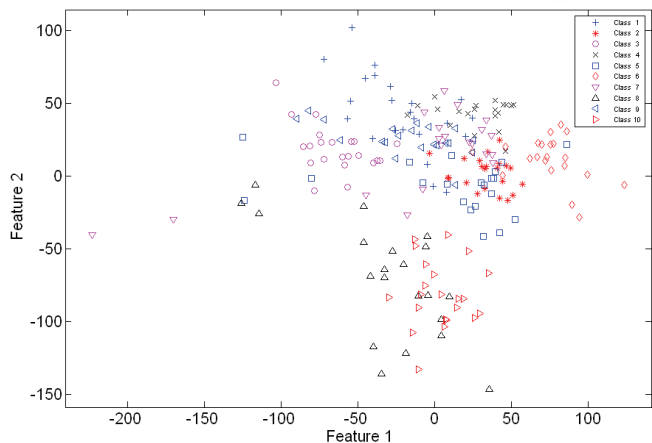


Fig. 9. Samples projected by Eigenfaces method.

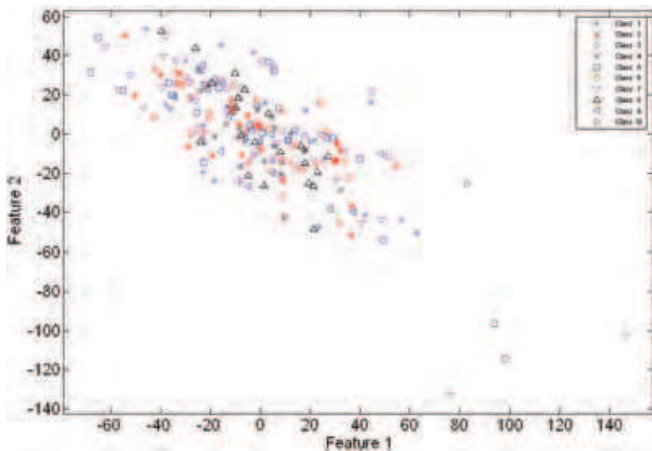


Fig. 10. Samples projected by Fisherfaces method.

Fig. 11 demonstrates the resultant data distribution of EigenExpression was worse than that of Eigenfaces (Fig. 9), but slightly worse than that of Fisherfaces (Fig. 10). The finding indicates that this method smears the classes together so that they are not linearly separable in the projected space. Note that our previous research (Tsai et al., 2005) (Tsai and Jan, 2005) also indicates that for an increase in the number of individuals the data distribution becomes more complex and intermixed.

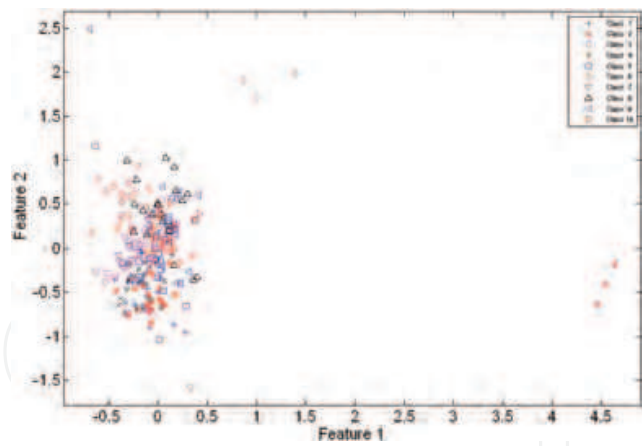


Fig. 11. Sample projected by EigenExpression method.

Fig. 12 demonstrates the resultant data distribution of FisherExpression method was less overlapped than those of Fisherfaces (Fig. 10) and EigenExpression methods (Fig. 11), but seems to be comparable to that of Eigenfaces (Fig. 9). This method promises to retain more discriminant information than that of Fisherfaces (Fig. 10). As a note of caution, it must be remembered that the resultant analyses from the above methods may be very application-specific.

The fused facial appearance and facial expression features of ten individuals were reduced to thirty features after using our method. The first two features are shown in Fig. 13. It shows that the fused features are well clustered after our method; it was far better than Eigenfaces, Fisherfaces, EigenExpression, and FisherExpression methods for classification. It indicates that the combined biometrics mutually contributed to the between-class separation.

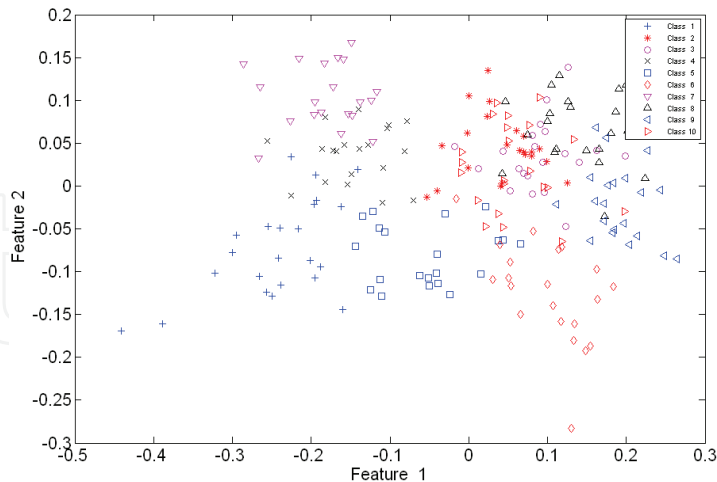


Fig. 12. Sample projected by FisherExpression method.

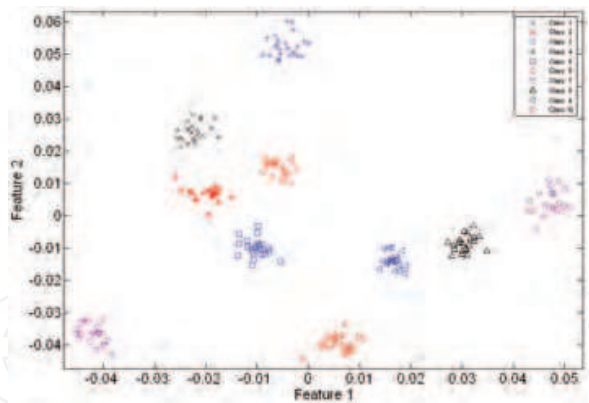


Fig. 13. Samples projected by our method of fused facial appearance and facial expression features.

9.3 Recognition Performance Comparison

9.3.1 Single Biometric Sources of Evidence: Face Versus Expressions

In this section, the assumption for the following experiments is that there is no significant difference in performance between using facial appearance or facial expressions as a biometric, given 1) use of the same PCA- and LDA-based algorithms for these two biometrics, and 2) use the different combinations for training and testing sets for experiment I and II. The recognition rates of these two experiments are computed based on the first 17 principle features. In experiment I, the baseline is the expression variation in the testing set. The experiment focuses on the recognition performance under only neutral images in the training set and facial expression images in the testing set. That is, all the neutral images were used for the training set and all the expression images were used for testing set.

Fig. 14 displays the recognition performance for experiment I. This performance is computed after 17 iterations. In terms of the facial appearance features, there was significant difference between EigenFace and FisherFace. The figure showed that the performance of EigenFace method was more stably increased than that of the FisherFace method after 17 iterations. In terms of the facial expression features, there was no significant difference between EigenExpression and FiserExpression methods. The recognition performances of these two methods were similar. Clearly, the findings indicate that if there exists facial expression changes, these changes degrade the recognition performances; as one would expect the more the facial expression changes vary from the neutral images, the worse the resulting recognition performances.

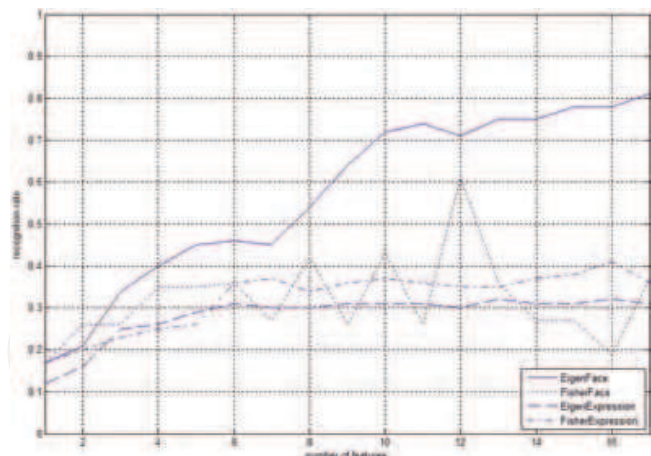


Fig. 14. Experiment I: Performance comparison - Face vs. Expressions (17 Iterations).

In experiment II, the baseline is for the only neutral faces in the testing set. The experiment focuses on the recognition performance under the condition that all the facial expression images are seen in the training stage and only neutral facial images in the testing set. The purpose here is to determine if by knowing all the individuals’ facial expression changes, could their neutral facial images be recognized during the testing stage. Fig. 15 displays the recognition performance for experiment II. This performance is computed after 17 iterations. In terms of the facial appearance features, the EigenFace method had better performance than that of the FisherFace, and the EigenFace method still showed stability in performance increase compared to that of the FisherFace method. In terms of the facial expression features, the EigenExpression method had better performance than that of the FisherExpression method, and the EigenExpression method showed stability in performance increase compared with that of the FisherExpression

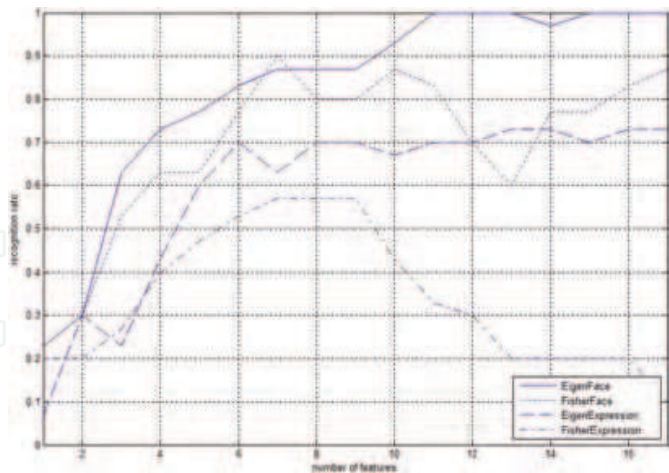


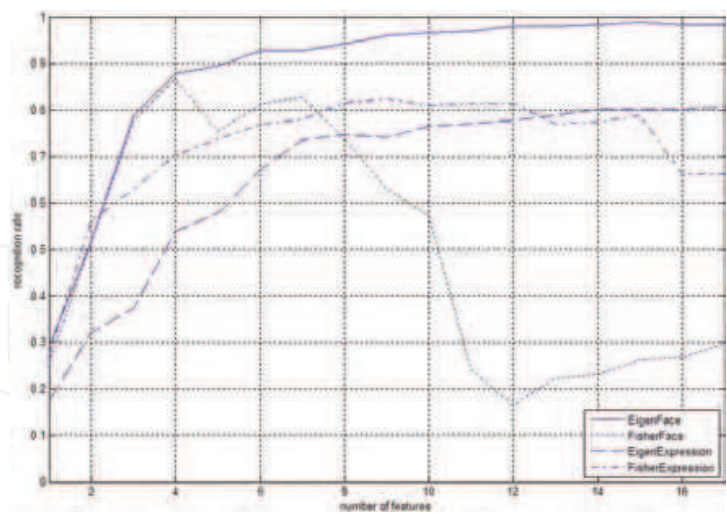
Fig. 15. Experiment II: Performance comparison - Face vs. Expressions (17 Iterations).

method. Clearly, the recognition performances of the four methods indicate the possibility that by learning individuals’ facial expression changes, one can test and recognize their individual neutral faces in the testing stage. Note that the LDA-based methods here show their limitation in performance stability due to the limited training data.

In experiment III, the baseline is all the facial expressions plus neutral faces have equal chances to be seen or used either in the training or testing set. The 10-fold cross-validation is used here for performance evaluation. In addition, we also use the leave-one-out cross-validation method so as to make use of the limited size of the database as the result of providing more consolidated performance evaluation for validating our hypothesis.

This performance is computed after 17 iterations. In Fig. 16(a), in terms of the facial appearance features, the EigenFace method had better performance than that of the FisherFace, and the FisherFace method in this figure showed its dramatic instability in recognition performance. In terms of the facial expression features, the performances of EigenExpression and FisherExpression methods were quite similar; however, the performance of the FisherExpression started to decrease after the 12 features. This figure showed the limitation of the LDA-based performance instability. Clearly, the recognition performances of the four methods showed that both facial appearance and expression features can be used for biometric recognition. Note that :

- (1) Fig. 15 and Fig. 16(a) look similar, and they indicate the importance of the facial expression changes seen in the training stage.
- (2) Fig. 16(b) shows similar performances to that of Fig. 16(a), except the performance of FisherFaces was increased stably. The limitation of the LDA-based method has been mitigated when the size of the training database is increased to the total number of images - 1 of the JAFFE database. This indicates the size of the JAFFE database is adequate and sufficient for validating our hypothesis in this thesis.
- (3) Overall, the results of these three experiments do not provide any significant evidence for rejecting the null hypothesis that facial appearance and facial of the facial appearance-based methods, one can see that these three experiments do show the importance of both facial appearance and facial expressions.



(a) 10-fold Cross Validation

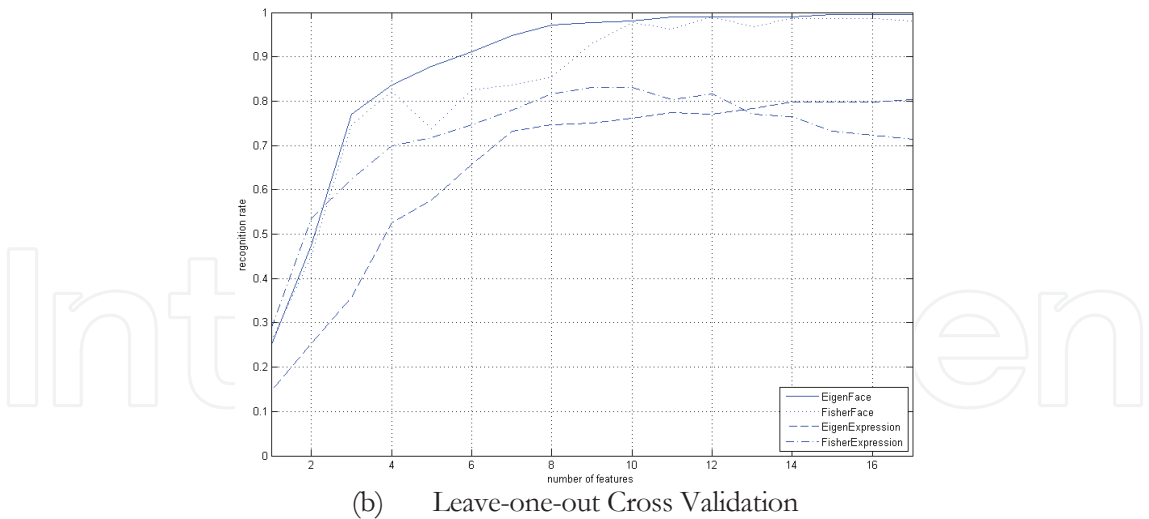


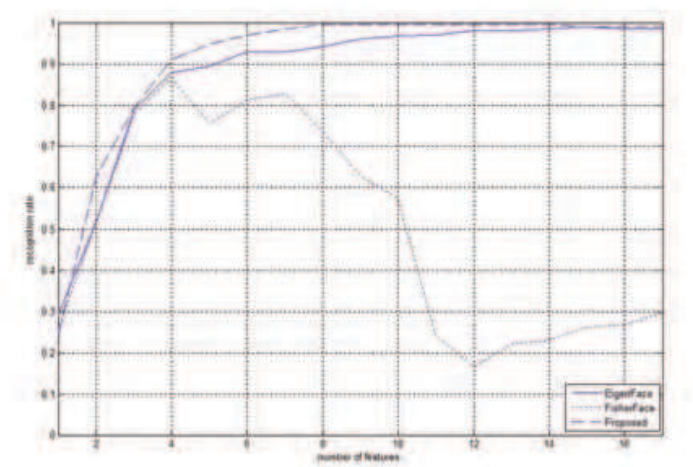
Fig. 16. Experiment III: Performance comparison - Face vs. Expressions (17 Iterations).

Especially, in Fig. 15 and Fig. 16, they indicate the importance of individuals’ facial expressions seen in the training set. Of course, there are several reasons that would affect the recognition performance results shown here; for example, inappropriate algorithms or dissimilar quality control conditions. In the following section, we are going to investigate the potential multimodal biometric system using facial appearance and facial expression features.

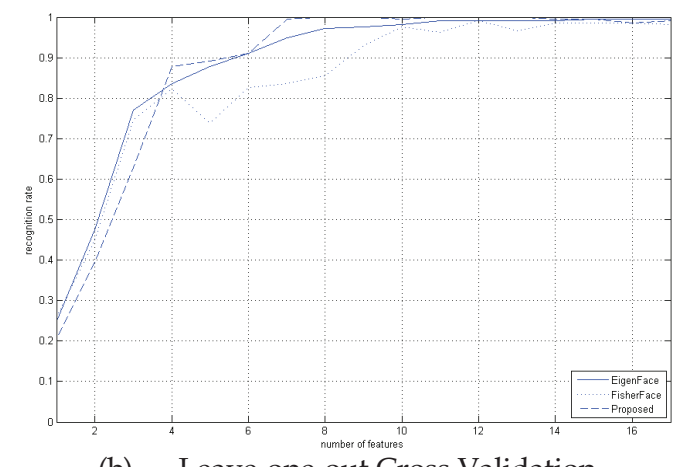
9.3.2 Multiple Biometric Sources of Evidence: Face Plus Expressions

A combination scheme shown in Fig. 4 is used to investigate the value of a multimodal biometric system using the facial appearance and facial expression features. In experiment IV, the baseline is the multiple biometric sources using facial appearance and facial expression features. The 10-fold cross validation is used here for performance validation. In addition, we also use the leave-one-out cross-validation method so as to make best use of the limited size of the database.

Fig. 17 and Fig. 18 display the recognition performances between the single and multiple modalities, respectively. Fig. 17(a) shows the comparison between face-based (single modality) and proposed (multiple modality) methods after 17 iterations for experiment IV. Our method using facial appearance and facial expression features had better recognition performance than that of either EigenFace or FisherFace method. Although the performances of the EigenFace and our methods were competitive, it was apparent that our method out-performed the EigenFace and FisherFace method when the dimensionality was very low (here 9).

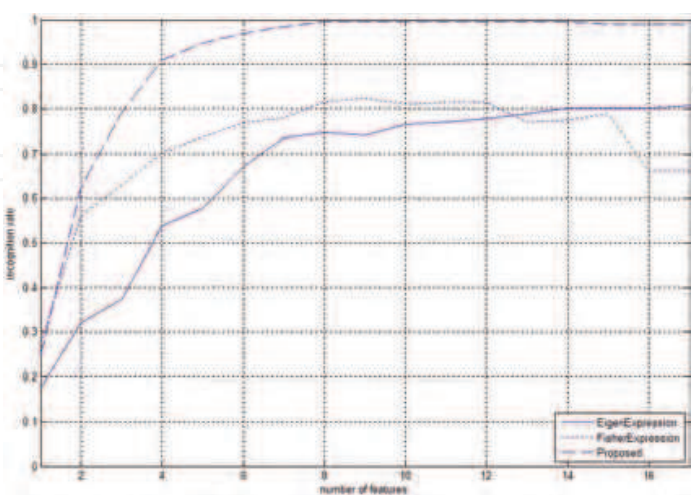


(a) 10-fold Cross Validation

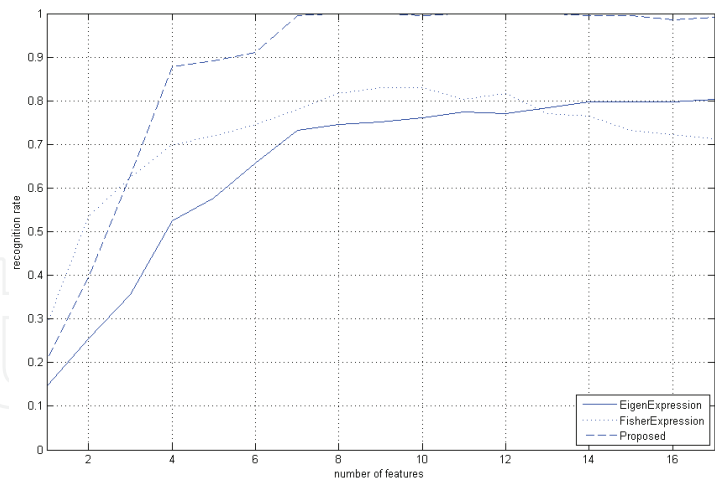


(b) Leave-one-out Cross Validation

Fig. 17. Experiment IV: The comparison between face-based and proposed methods (17 Iterations).



(a) 10-fold Cross Validation



(b) Leave-one out Cross Validation

Fig. 18. Experiment IV: The comparison between expression-based and proposed methods (17 Iterations).

Fig. 18(a) shows the comparison between expression-based (single modality) and proposed (multiple modality) methods after 17 iterations for the experiment IV. Our method using facial appearance and facial expression features had better recognition performance than that of either EigenExpression or FisherExpression method. Our method again outperformed the single modality methods when the dimensionality was very low (here 9). These two figures clearly indicate the possibility to combine the facial appearance and facial expression features for multiple biometric sources. Fig. 17(b) and Fig. 18(b) showed similar performance to those in Fig. 17(a) and Fig. 18(a), except the performance of FisherFaces was increased stably. The limitation of the LDA-based method has been mitigated when the size of the training database is increased to the total number of images – 1 of the JAFFE database. This indicates the size of the JAFFE database is adequate and sufficient for validating our hypothesis in this thesis.

9.3.3 Confidence Analysis of Prediction Performance

According to [75] , the CI metric is used for measuring how confident the estimated experimental results from experiment IV (using 10-fold and leave-one-out cross-validations) are close to the true prediction performance. In this section, we will conduct the Confidence Interval measurement of the Prediction Performance so as to consolidate and to draw our validating conclusion to prove our assumption that intra-personal variation of facial behavior can assist the separation of extra-personal separation. The CI metric is applied to analyze the confidence degree of the two performance evaluations on evaluating our hypothesis; namely, the 10-fold cross-validation and leave-one-out cross-validation methods. We will only focus on analyzing the experimental results of comparative methods when the number of features is nine ($d = 9$)⁴.

⁴ In the literature, it is recommended the optimal number of principle components is the number of classes minus one.

Fig. 19 and Fig. 20 show the confidence interval analyses for the experimental results obtained from the 10-fold cross-validation and leave-one-out cross-validation.

In Fig. 19, there are wide ranges of confidence intervals for different comparative methods on different confidence values. For example, if we increase our confidence values from 20% to 99.8% for EigenFaces, we can see that its confidence boundaries are becoming wider and wider. Also, we can see that other methods have similar outcomes. The CI intervals tells us that the estimated experimental results obtained from 10-fold cross-validation have wide variations from the actual true experimental results. However, even though 10-fold cross-validations gave us conservative experimental results, we can still see that our method has a narrower range than any of the other methods (Please SEE Table 1). Clearly, it indicates to us that we can trust the estimated experimental results in Experiment IV. It also shows us that our method has the advantage over other comparative methods. Our result in this figure is an indication of the potential of using facial behavior combined with the facial appearance for personal identification. For further substantiation, we also use the CI metric for analyzing the leave-one-out cross-validation.

In Fig. 20, there are small ranges of confidence intervals for different comparative methods on different confidence values. We can see that after using the leave-one-out cross-validation, the ranges of confidence intervals are narrower than those in Fig. 19. Even though the confidence values are increased from 20% to 99.8%, the ranges of CI for different comparative methods are still small. This figure tells us that by using the leave-one-out cross-validation, our limited database has been maximized. In particular, we can see from the figure that our method has the smallest ranges of CI for different confidence values among other comparative methods (Please see Table 1. Estimated success rates of comparative methods and their corresponding confidence interval boundaries ($c = 99.8\%$) using 10-fold cross-validation.

).

To sum up, the CI analyses for both 10-fold cross-validation and leave-one-out cross-validation conclude that our method validate our hypothesis that the intra-personal variations of facial behavior may assist the extra-personal separations.

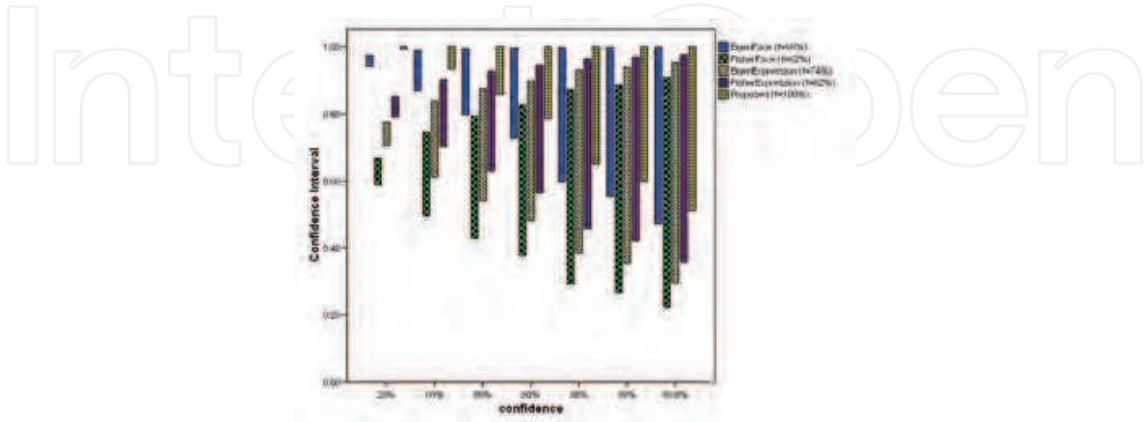


Fig. 19. Confidence Interval for the experiment IV results using 10-fold cross-validation ($N = 10 ; d = 9$).

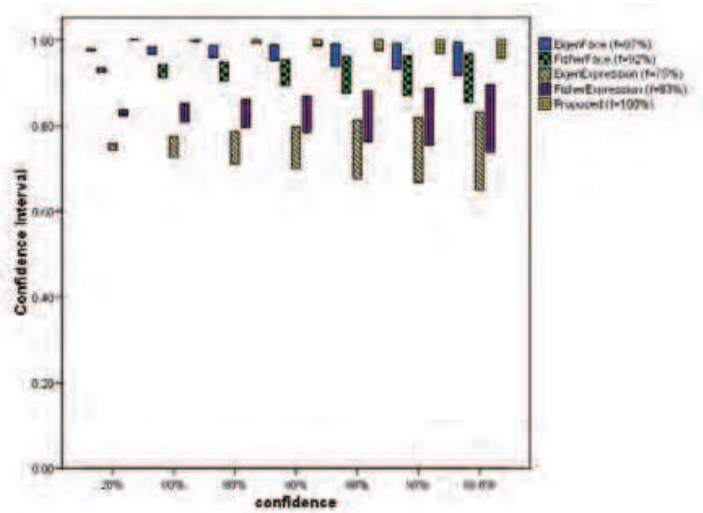


Fig. 20. Confidence Interval for the experiment IV results using leave-one-out cross-validation ($N = 213 ; d = 9$).

Methods	f	p^-	p^+
EigenFaces	0.96	0.472185	0.998449
FisherFaces	0.629	0.222549	0.909433
EigenExpression	0.742	0.29596	0.951634
FisherExpression	0.824	0.356099	0.975391
Our Method	1	0.511559	1

Table 1. Estimated success rates of comparative methods and their corresponding confidence interval boundaries ($c = 99.8\%$) using 10-fold cross-validation.

Methods	f	p^-	p^+
EigenFaces	0.976526	0.918645	0.993517
FisherFaces	0.929577	0.855038	0.967257
EigenExpression	0.751174	0.650202	0.830593
FisherExpression	0.830986	0.737872	0.895699
Our Method	1	0.957096	1

Table 2. Estimated success rates of comparative methods and their corresponding confidence interval boundaries ($c = 99.8\%$) using leave-one-out cross-validation.

9.4 Discussion

Yan (Yang, 2002) states that the PCA method aims to extract the most expressive bases in which the most optimal eigen-vectors are extracted in order to reduce the reconstruction error to the minimum. However, some unwanted variations due to illumination, poses, and/or facial expressions may be caught in the sense of this correlation-dependency method during the extraction stage. That is, these variations may increase the reconstruction error. Therefore, this method may be optimal for obtaining a good PCA projection in a correlation sense, whereas, it may be less optimal in the sense of the classification viewpoint.

Nonetheless, it is interesting to note that our experimental results shown in Fig. 17 seem to contradict the good empirical results in (Belhumeur et al., 1997). They state that “although PCA achieves larger total scatter, FLD achieves greater between-class scatter, and, consequently, classification is simplified (p.714)”. In fact, the Eigenfaces method under various facial expression changes in our experiments shows its high recognition rates. It is likely that the variation of facial expressions caught in this extraction stage also provides distinctive characteristics about individuals, and PCA achieves larger total scatter; hence, yields more spaces and less overlap in data distribution. Other possible reason that causes this discrepancy from theirs is possibly because they use leave-one-out strategy to evaluate their system based on the small database⁵, and it contains different variations in illumination, occlusion, and facial expressions⁶. These variations may lead to degeneration to the classification performance when applying this second-order correlated statistical method.

It has been suggested in (Adini et al., 1997) that the intra-personal variations due to illumination and viewing direction are almost always larger than the extra-personal differences. Other papers (Belhumeur et al., 1997) also address the variation due to facial expressions observes a similar effect in face identity. Therefore, it is debatable that the variation due to the internal facial changing factors could be considered as noise that affect the facial recognition or be considered as another biometric cue that could be used solely or assist in facial recognition. Note that we consider the variations due to such things as illumination, poses, or occlusions as external factors of a face and the variations due to facial expressions as internal factors of a face.

It is surprising to see from Fig. 13, that after the fusion of the facial appearance and expression features, the data distribution has become more compact and non-overlapped among individuals, in comparison with other methods as shown in Fig. 9 through Fig. 12. It seems to indicate that the internal changing factors may lead us to another interesting research area.

Moreover, the findings from Fig. 9 and Fig. 11 also demonstrate that each dimension from the single biometric may contribute some mutually discriminate information in fusion, even though the data distributions of these single methods are overlapped.

To sum up, our framework has improved recognition performance of each of facial appearance-based and facial expression-based modality. The findings also indicate the potential of using facial expression behavior for either single or multiple biometrics recognition. In particular, our method achieves almost 100% recognition rate using only 9 features.

10. Conclusion and Future Work

This chapter assessed the possibility of using facial behaviour as another individual trait for personal identification and identification improvement. Facial expression variations were previously thought of as noise that would degrade the classification performance, so

⁵ They use sixteen subjects, and each has 10 images.

⁶ The facial expressions include happy, sad, winking, sleepy, and surprised. Here, those variations are considered noise that degrades classification performance.

researchers tried to build a robust facial recognition system insensitive to facial expression changes. We took the opposite view and assumed that the dynamic information of intra-personal facial behaviour was useful not only as another behavioural biometric but could also assist the extra-personal separation for improvement of recognition performance. This innovative approach is motivated by psychophysiology regarding the way humans recognize one another from either their facial appearance or facial behaviour or both. Our findings support this hypothesis. Experiment results demonstrated that our approach outperformed other conventional methods, with highest f-value and p-values. To sum up, the research presented in this chapter opens up a new research direction that can apply to either single biometrics or multiple biometrics. More work needs to be done in this new research direction, such as implementing a larger ground truth database with spatio-temporal information of dynamic facial motion of expression for 3D modeling.

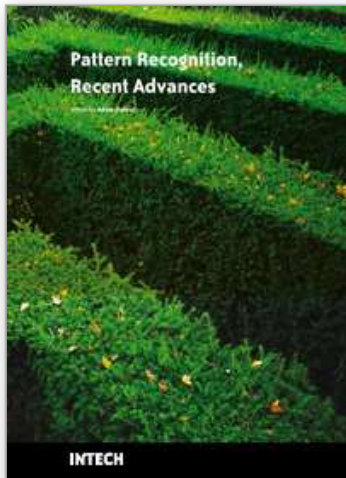
11. References

- ADINI, Y., MOSES, Y. & ULLMAN, S. (1997) Face recognition: the problem of compensating for changes in illumination direction. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19, 721-732.
- BELHUMEUR, P. N., HESPANHA, J. P. & KRIEGMAN, D. J. (1997) Eigenfaces versus Fisherfaces: Recognition using class specific linear projection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19, 711-729.
- BISHOP, C. M. (1995) *Neural Networks for Pattern Recognition*, Oxford, Oxford University Press.
- BOYER, K. W., GOVINDARAJU, V. & RATHA, N. K. (2007) Introduction to the Special Issue on Recent Advances in Biometric Systems [Guest Editorial]. *IEEE Transactions on Systems, Man, and Cybernetics, Part B* 37, 1091-1095.
- BRONSTEIN, A. M., BRONSTEIN, M. M. & KIMMEL, R. (2003) Expression-Invariant 3D Face Recognition. *Lecture Notes in Computer Science*.
- BRONSTEIN, A. M., BRONSTEIN, M. M. & KIMMEL, R. (2007) Expression-Invariant Representations of Faces. *IEEE Transactions on Image Processing*, 16, 188-197.
- BRUNELLI, R. & FALAVIGNA, D. (1995) Person identification using multiple cues. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17, 955-966.
- CHANG, K., BOWYER, K. W., SARKAR, S. & VICTOR, B. (2003) Comparison and combination of ear and face images in appearance-based biometrics *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25, 1160-1165.
- CHANG, K. I., BOWYER, K. W. & FLYNN, P. J. (2005) An Evaluation of Multimodal 2D+3D Face Biometrics. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27, 619 - 624.
- CHANG, K. I., BOWYER, K. W., FLYNN, P. J. & CHEN, X. (2004) Multi-biometrics using facial appearance, shape and temperature. *IEEE International Conference on Automatic Face and Gesture Recognition*.
- CHEN, L. F., LIAO, H. Y. M., KO, M. T., LIN, J. C. & YU, G. J. (2000) A new LDA-based face recognition system which can solve the small sample size problem. *Pattern Recognition*, 33, 1713-1726.
- DUDA, R. O., HART, P. E. & STORK, D. G. (2000a) *Pattern Classification*, New York, Wiley

- DUDA, R. O., HART, P. E. & STORK, D. G. (Eds.) (2000b) *Pattern Classification*, New York, Wiley.
- FAUNDEZ-ZANUY, M. (2005) Data fusion in biometrics. *IEEE Aerospace and Electronic Systems Magazine*. Phoenix, AZ, USA.
- FR BA, B., ROTHE, C. & K BLBECK, C. (2000) Statistical Sensor Calibration for Fusion of Different Classifiers in a Biometric Person Recognition Framework. *Multiple Classifier Systems*.
- HONG, L., JAIN, A. & PANKANTI, S. (1999) Can Multibiometrics Improve performance? *Proceedings AutoID'99, Summit*. NJ.
- HU, W., TAN, T., WANG, L. & MAYBANK, S. (2004) A survey on visual surveillance of object motion and behaviors. *IEEE Transaction on Systems, Man, and Cyernetics - Part C: Applications and Reviews*, 34, 334-352.
- JAIN, A. K., DASS, S. C. & NANDAUMAR, K. (2004a) Can Soft Biometric Traits Assist User Recognition? *Proceedings of SPIE on Biometric Technology for Human Identification*. Orlando, Florida.
- JAIN, A. K., NANDAKUMAR, K., ULUDAG, U. & LU, X. (2006) Multimodal Biometrics: Augmenting Face With Other Cues. IN ZHAO, W. & CHELLAPPA, R. (Eds.) *Face Processing: Advanced Modelling and Methods*. New York, Elsevier.
- JAIN, A. K., ROSS, A. & PRABHAKAR, S. (2004b) An introduction to biometric recognition. *IEEE Transactions on Circuits and Systems for Video Technology*, 14, 4-20.
- JUWEI, L., PLATANOTIS, K. N. & VENETSANOPOULOS, A. N. (2003) Regularized D-LDA for face recognition.
- KIRBY, M. & SIROVICH, L. (1990) Application of the Karhunen-Loeve procedure for the characterization of human faces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12, 103-108.
- KONG, H., WANG, L., TEOH, E. K., WANG, J. G. & VENKATESWARLU, R. (2005) A framework of 2D Fisher discriminant analysis: application to face recognition with small number of training samples.
- LIN, H. & ANIL, J. (1998) Integrating faces and fingerprints for personal identification. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20, 1295-1307.
- LIU, Y., SCHMIDT, K. L., COHN, J. F. & MITRA, S. (2003) Facial asymmetry quantification for expression invariant human identification. *Computer Vision and Image Understanding: CVIU*, 91, 138-159.
- LIU, Y., SCHMIDT, K. L., COHN, J. F. & WEAVER, R. L. (2002) Facial asymmetry quantification for expression invariant human. *Proceedings of Fifth IEEE International Conference on Automatic Face and Gesture Recognition*.
- LIU, Z. & SARKAR, S. (2007) Outdoor recognition at a distance by fusing gait and face. *Image and Vision Computing*, 25, 817-832.
- LU, J., PLATANOTIS, K. N. & VENETSANOPOULOS, A. N. (2003a) Face recognition using kernel direct discriminant analysis algorithms. *IEEE Transactions on Neural Networks*, 14, 117-126.
- LU, J., PLATANOTIS, K. N. & VENETSANOPOULOS, A. N. (2003b) Face recognition using LDA-based algorithms. *IEEE Transactions on Neural Networks*, 14, 195-200.
- LU, J., PLATANOTIS, K. N. & VENETSANOPOULOS, A. N. (2003c) Regularized discriminant analysis for the small sample size problem in face recognition. *Pattern Recognition Letters*, 24, 3079-3087.

- LU, J., PLATANIOTIS, K. N. & VENETSANOPOULOS, A. N. (2005) Regularization studies of linear discriminant analysis in small sample size scenarios with application to face recognition. *Pattern Recognition Letters*, 26, 181-191.
- MARTINEZ, A. M. (2000) Semantic access of frontal face images: the expression-invariant problem. *Proceedings of IEEE Workshop on Content-based Access of Image and Video Libraries*.
- PHILLIPS, P. J., HYEONJOON, M., RIZVI, S. A. & RAUSS, P. J. (2000) The FERET evaluation methodology for face-recognition algorithms. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22, 1090-1104.
- ROSS, A. & JAIN, A. K. (2003) Information Fusion in Biometrics. *Pattern Recognition Letters*, 24, 2115-2125.
- SHAKHNAROVICH, G., LEE, L. & DARRELL, T. (2001) Integrated face and gait recognition from multiple views. *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*.
- TOH, K.-A., KIM, Y., LEE, S. & KIM, J. (2008) Fusion of visual and infra-red face scores by weighted power series. *Pattern Recognition Letters*, 29, 603-615.
- TSAI, P. & JAN, T. (2005) Expression-Invariant Face Recognition System Using Subspace Model Analysis. *Proceedings of IEEE International Conference on Systems, Man and Cybernetics*, Hawaii, USA.
- TSAI, P., JAN, T. & HINTZ, T. (2005) Expression-Invariant Face Recognition for Small Class Problem. *Proceedings of IEEE International Conference on Computational Intelligence for Measurement Systems and Applications*. Sicily, Italy.
- WECHSLER, H. (2007) *Reliable Face Recognition Methods: System Design, Implementation and Evaluation*, Springer.
- YANG, M.-H. (2002) Kernel Eigenfaces vs. Kernel Fisherfaces: Face Recognition Using Kernel Methods. *Proceedings of the Fifth IEEE International Conference on Automatic Face and Gesture Recognition*.
- ZHOU, X. & BHANU, B. (2007) Integrating Face and Gait for Human Recognition at a Distance in Video. *IEEE Transactions on Systems, Man, and Cybernetics, Part B*, 37, 1119-1137.

IntechOpen



Pattern Recognition Recent Advances

Edited by Adam Herout

ISBN 978-953-7619-90-9

Hard cover, 524 pages

Publisher InTech

Published online 01, February, 2010

Published in print edition February, 2010

Nos aute magna at aute doloreetum erostrud eugiam zzriuscipsum dolorper iliquate velit ad magna feugiamet, quat lore dolore modolor ipsum vullutat lorper sim inci blan vent utet, vero er sequatum delit lortion sequip eliquatet ilit aliquip eui blam, vel estrud modolor irit nostinc iliquiscinit er sum vero odip eros numsandre dolessisisim dolorem volupta tionsequam, sequamet, sequis nonnulla conulla feugiam euis ad tat. Igna feugiam et ametuercil enim dolore commy numsandiam, sed te con hendit iuscidunt wis nonse volenis molorer suscip er illan essit ea feugue do dunt utetum vercili quamcon ver sequat utem zzriure modiat. Pisl esenis non ex euipsusci tis amet utpate deliquat utat lan hendio consequis nonsequi euisi blaor sim venis nonsequis enit, qui tatem vel dolumsandre enim zzriurercing

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Pohsiang Tsai, Tich Phuoc Tran and Longbing Cao (2010). A New Multimodal Biometric for Personal Identification, Pattern Recognition Recent Advances, Adam Herout (Ed.), ISBN: 978-953-7619-90-9, InTech, Available from: <http://www.intechopen.com/books/pattern-recognition-recent-advances/a-new-multimodal-biometric-for-personal-identification>

INTECH
open science | open minds

InTech Europe

University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China

Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

© 2010 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the [Creative Commons Attribution-NonCommercial-ShareAlike-3.0 License](https://creativecommons.org/licenses/by-nc-sa/3.0/), which permits use, distribution and reproduction for non-commercial purposes, provided the original is properly cited and derivative works building on this content are distributed under the same license.

IntechOpen

IntechOpen