Proceedings of the 2009 IEEE International Conference on Systems, Man, and Cybernetics
San Antonio, TX, USA - October 2009

Two-View Face Recognition Using Bayesian Fusion

Grace Shin-Yee Tsai
Department of Electrical Engineering
National Central University
Jhongli City, Taoyuan County 32001, Taiwan
graceshinyeetsai@gmail.com

Angela Chih-Wei Tang
Department of Communication Engineering
National Central University
Jhongli City, Taoyuan County 32001, Taiwan
cwtang@ce.ncu.edu.tw

***Abstract*—Two-dimensional face recognition suffered from pose changes, while three-dimensional approaches are with high computational complexity. Motivated by this, a two-view face recognition system for digital home is presented in this paper. Besides the improvement in recognition rate, this system reduces the misclassification that could occur in traditional single-view systems. The proposed system fuses the individual recognition results of two images of the same identity with different viewing angles based on Bayesian theory. Bayesian approach uses the similarity of each person and is trained by determining the reliability of each identity of the two channels. A frontal view and a side view are chosen since they convey the most important information of human faces. Each input image is sent into its corresponding channel to obtain a 2D face recognition result. Within each channel, PCA and SVM are applied. Different form traditional PCA based approaches, SVM classifiers are used instead of minimum distance classifier to enhance the robustness. Our experimental results show that this two-view face recognition system has achieved a higher recognition rate compared with traditional 2D single-view face recognition systems.**

***Keywords*—PCA, SVM, Bayesian theory, fusion, face recognition, multi-channel.**

I. INTRODUCTION

Examples of face recognition include information security, access control, and human computer interaction. In the past, plenty of algorithms have been proposed to improve the recognition rate and its practicability in real-time applications. However, face recognition are still facing two major challenges, pose and illumination variations [1].

2D face recognition systems are extremely sensitive to pose variations for its strong dependency on full still input images. Thus, in recent years, many 3D based face recognition systems have been proposed. 3D face recognition methods include stereo reconstruction [2], SFS (shape from shading) [3], range-based reconstruction [4], etc. However, these methods usually involved with a large scale of computations or required some manual initializations.

Another new face recognition approach that was proposed is multi-channel face recognition systems. These systems are mostly based on 2D face analysis techniques and are built to enhance the robustness against posture or illumination variations. There are several approaches of multi-channel face recognition systems, including recognition from multiple still images [2] [5], multi-model face recognition [6] and multi-algorithm face recognition [7] [8]. The inputs of these algorithms may acquire a single image or multiple images. For those with single input images, a classifier is added before face recognition to determine which channel is the most suited face recognition for the current input. For those with multiple input images, a fusion mechanics is needed to combine the results from each channel. It is worth notice that face recognition systems with multiple input images greatly reduced the confusions of identities that occur in those with single input. That is, the errors occurred in single-input systems are mainly caused by identifying a known person as someone unknown. In many real-world applications, such as human computer interactions, misclassifications will cause a lot of troubles.

Motivated by the reasons above, this paper proposes a two-view face recognition system. Two still images with different viewing angles of the same identity are sent into different channels. Within each channel, a holistic 2D analysis on face images is performed. A combined method of PCA and SVM is applied in the system [8] [9]. To fuse the two channels, the decision mechanics, based on Bayesian theory, designed in [10] is adopted. In recent multi-view works in [2] [5], AdaBoost are applied to its selection mechanics. This approach considers the reliability of the channels which considers several features, such as wavelet features, face recognition matching distance features, consecutive time features. In each channel/view, a reliability function for a channel is generated by training the AdaBoost classifier with images from all identities. At the recognition stage, the selector directly outputs the result from the channel with the higher reliability. However, the recognition accuracy varies with identities and viewing angles. Therefore, in our adopted Bayesian approach, the decision mechanics is trained by considering the relations of reliability between two identities within each viewing angle. Parameters of the Bayesian fusion are the probability of each person being recognized as each of the possible identity. Miss identifications between certain people occurred in a channel might appear correct classification in the other channel. Through the Bayesian decision mechanics, these miss identifications will reduce sufficiently. In other words, the two channels are somehow acting as a compensator of each other by using Bayesian fusion.

This paper is organized as followed. Section 2 reviews the basic concepts of principal component analysis. Section 3 gives a brief introduction to SVM classification. Section 4 describes the proposed face recognition scheme and shows how the Bayesian based decision mechanics are designed and trained. Section 5 provides experimental results of the two-channel

978-1-4244-2794-9/09/$25.00 ©2009 IEEE

systems and its comparison to traditional approaches. Finally, we give conclusions in Section 6.

II. PRINCIPAL COMPONENT ANALYSIS

In previous works, there have been several projection methods such as PCA (Principal Component Analysis or Eigenfaces) and LDA (Linear Discriminant Analysis or Fisherfaces) employed for face recognition. PCA can be used for dimensionality reduction in a data set. LDA is similar to PCA in looking for linear combinations of variables which best characterize the data. LDA will outperform PCA if a large database is adopted or higher dimensional subspaces are selected [11] [12]. In our proposed face recognition for digital home, the total number of people is reasonably limited in the database. In this case, PCA is preferred since less images of each person is needed and less dimensionality of the features is used for classification. Thus, we adopt PCA in our system and the concept of PCA will be described as follows. More details could be found in [13].

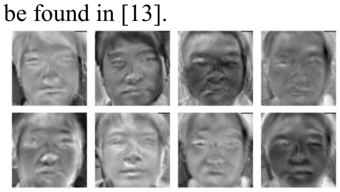

Figure 1. Examples of eigenfaces.

Given an N-dimensional vector representation of each face image, PCA finds a feature space whose basic vectors correspond to the directions of the largest variance among the training set in the original space. Let $\mathbf{W}$ represents the linear transformation that projects the original N-dimensional space onto the k-dimension feature space where k is usually much smaller than N. The new feature vectors $y_i \in \Re^f$ are determined by $y_i = \mathbf{W}^T x_i$, for $i = 1, \ldots, M$, where M is the total number of training data. Let matrix $\mathbf{X}$ represents the face images of the training set where $\mathbf{X} = [\mathbf{x_1}, \mathbf{x_2}, \mathbf{x_3} \ldots \mathbf{x_M}]$and $\mathbf{X}$ is of dimension$N \times M$, where N is the number of pixels in the face image. The matrix $\mathbf{X}'$ is the difference between face images and the average face image $\bar{\mathbf{x}}$:

$$\mathbf{X}' = [(\mathbf{x_1} - \bar{\mathbf{x}}), (\mathbf{x_2} - \bar{\mathbf{x}}), (\mathbf{x_3} - \bar{\mathbf{x}}), \ldots, (\mathbf{x_M} - \bar{\mathbf{x}})], \quad (1)$$

where

$$\bar{\mathbf{x}} = \frac{1}{M} \sum_{i=1}^{M} \mathbf{x_i}. \quad (2)$$

PCA seeks a set of $M - 1$ orthogonal vectors, e_i, which best describe the distribution of the data. The typical method of computing the principal components is to solve the eigen-problem of the covariance matrix $\mathbf{C} = \mathbf{X}'\mathbf{X}'^{\mathbf{T}}$, an $N \times N$ matrix. This will normally be a huge matrix, and a full computation of the eigenvectors is impractical. Alternatively, the eigenvectors can be computed by using an $M \times M$ eigenvector calculation since there are only $M - 1$ nonzero eigenvalues. The relation between the two methods can be determined easily [14].

The eigenvectors e_i and eigenvalues λ_i of $\mathbf{C}$, where are related to the eigenvectors $\widehat{e_i}$ and eigenvalues μ_i of the $M \times M$ matrix

$$\mathbf{D} = \mathbf{X}'^{\mathbf{T}}\mathbf{X}'. \quad (3)$$

By computing, the eigenvectors e_i and eigenvalues λ_i of $\mathbf{C}$ can be determined by [15]

$$e_i = (\mathbf{X}'\widehat{e_i}), \quad \lambda_i = \mu_i. \quad (4)$$

With this analysis, the calculations are greatly reduced from the order N, the number of pixels, to the order, M, the number of the images in the training set.

Once the eigenvectors e_i and eigenvalues λ_i of $\mathbf{C}$ are found, the face space y_i could be defined. Since eigenvectors appear like faces, as shown in Fig 1, they are called eigenfaces. These eigenfaces are the basic directions of the face space y_i. In other words, the original face images could be reconstructed by using merely the linear combination of these eigenfaces. Every eigenface corresponds to an eigenvalue. A larger eigenvalue means the eigenvector captures a larger amount of variation among the training set. Thus, we could choose only the desired k major eigenvectors to be the orthogonal vectors for the face space y_i. [11] The number of eigenvectors is typically chosen by

$$\frac{\text{Sum of the k major eigenvalues}}{\text{Sum of all eigenvalues}} = \frac{\sum_{i=1}^{k} \lambda_i}{\sum_{i=1}^{M-1} \lambda_i} > \vartheta, \quad (5)$$

where ϑ represents the similarity of the actual face images and the reconstructed face images.

Using the k eigenfaces, the linear transformation, $\mathbf{W} = [w_1, w_2, w_3, \ldots, w_k]$, between the original and face space could be determined by simply projecting the face images onto each eigenface,

$$w_i = e_i^T(\mathbf{x_i} - \bar{\mathbf{x}}), \text{ for } i = 1, \ldots, M. \quad (6)$$

Each w_i represents the weighting of the corresponding eigenface in the linear combination of the reconstructed face. Thus, the transformation matrix $\mathbf{W}$ could be the features of each face images which reduces the dimension of each data from N to k. Therefore, face images at the testing stage of face recognition could be transformed into the face space by using (6) and obtain a transformation matrix.

III. SUPPORT VECTOR MACHINES

In traditional PCA based approaches, recognitions are performed using minimum distance classification on the face space. These systems are very sensitive to illumination and pose variations due to a lack of statistical analysis. On the other hand, face recognition using SVM classification are said to be more reliable [16]. However, SVM based approaches usually lead to a larger amount of computations due to the huge data size of face images. Thus, feature extraction is necessary before applying SVM classification. By combining PCA and SVM, PCA acted as a feature extraction method that would greatly reduce the dimensionality of the face pattern for SVM classification. SVM provides a more reliable classification methods for PCA based approaches. Therefore, we adopt the combination of PCA and SVM in the face recognition system.

Support vector machines (SVMs) are a set of supervised learning techniques for data classification and regression [16] [17]. SVMs have been shown to give good generalized performance in various recognition related applications.

Typically, a classification task of SVM involved with training and testing set of data instances. Each instance in the training set contains a class label and a set of feature data. At the training stage, SVM is trained to determine a model for classification at testing stage where only the feature data are given. In the following, the basic concept of a binary SVM classification and its extensions to multiple-class classification implements will be introduced [18] [19].

A. Basic Concept (Binary Classification)

Let us consider the simplest case of SVM, the linear form. Consider a binary classification task with a training set of k examples by their feature-label pairs $(\mathbf{x}_\mathrm{i}, \mathrm{y}_\mathrm{i})$, for $\mathrm{i} = 1, 2, \ldots, \mathrm{k}$. $\mathbf{x}_\mathrm{i}$ denotes the feature data and y_i represents the class label where $\mathrm{y}_\mathrm{i} \in \{1, -1\}$. Suppose we have several hyperplanes, or decision boundaries, that separate a set of positive examples from the set of negative examples. A point $\mathbf{x}$ that lies on the decision boundary satisfies

$$\mathbf{w}^\mathrm{T}\mathbf{x} + \mathrm{b} = 0\ , \tag{7}$$

where $\mathbf{w}$ denotes a vector orthogonal to the decision line and $\mathrm{b}/\|\mathbf{w}\|$ is the perpendicular distance between the line and the origin. Define

$$\mathbf{x}^+ = \mathbf{w}^\mathrm{T}\mathbf{x} + \mathrm{b} = +1 \tag{8}$$

as the positive plan containing the closest positive example and

$$\mathbf{x}^- = \mathbf{w}^\mathrm{T}\mathbf{x} + \mathrm{b} = -1 \tag{9}$$

as the negative plane with the nearest negative example. The distance between the positive and negative plane is called the margin of the decision line. The margin $\mathbf{m}$ could be determined by

$$\mathbf{m} = |\mathbf{x}^+ - \mathbf{x}^-| = \frac{1-\mathrm{b}}{\|\mathbf{w}\|} - \frac{-1-\mathrm{b}}{\|\mathbf{w}\|} = \frac{2}{\|\mathrm{W}\|}. \tag{10}$$

For the case of linearly separable data, support vector algorithm seeks the decision hyperplane that has the largest margin with all training data satisfying the constraint

$$(\mathbf{w}^\mathrm{T}\mathbf{x}_\mathrm{i} + \mathrm{b})\mathrm{y}_\mathrm{i} \geq 1 \text{ for } \mathrm{i} = 1, 2, \ldots, \mathrm{k}. \tag{11}$$

Thus, we are now working on an optimization problem such that

$$\min_{\mathrm{w,b}} \frac{1}{2}\mathbf{w}^\mathrm{T}\mathbf{w} \tag{12}$$

which satisfies the constraint of (11) by noting that the magnitude of $\mathbf{m}$ merely scales $\mathbf{w}$ and b, and does not change the decision boundary.

However, not all training sets are linearly separable. While dealing with non-linearly separable problems, we allow an error ϵ_i in classification for some examples. The error is defined as the distance between the example and the desired line (positive or negative line). Hence, besides maximizing the margin, we also want to minimize the total error. Accordingly, the optimization problem could be modified into a more general form as follows

$$\min_{\mathrm{w,b}} \frac{1}{2}\mathbf{w}^\mathrm{T}\mathbf{w} + \mathrm{C}\sum_{\mathrm{i}=1}^{\mathrm{k}} \epsilon_\mathrm{i}, \tag{13}$$

subject to constraints

$$(\mathbf{w}^\mathrm{T}\mathbf{x}_\mathbf{i} + \mathrm{b})\mathrm{y}_\mathrm{i} \geq 1 - \epsilon_\mathrm{i} \text{ and } \epsilon_\mathrm{i} \geq 0, \text{ for } \forall \mathrm{i}. \tag{14}$$

In order to solve this optimization problem, the Lagrange multiplier methods are introduced. By Lagrange multiplier, every training example has a corresponding Lagrange multiplier α. Once we have all the Lagrange multipliers, the vector $\mathbf{w}$ could be reconstructed as a weighted combination of the training examples by

$$\mathbf{w} = \sum_{\mathrm{i}=1}^{\mathrm{k}} \alpha_\mathrm{i}\mathrm{y}_\mathrm{i}\,\mathbf{x}_\mathbf{i} \tag{15}$$

and

$$\mathrm{b} = \mathbf{w}^\mathrm{T}\mathbf{x}_\mathbf{i} - \mathrm{y}_\mathrm{i} \text{ for some } 0 < \alpha_\mathrm{i} < C\ . \tag{16}$$

Note that for $0 < \alpha_\mathrm{i} < C$, the corresponded training examples of these α_i are called support vectors (SV). However, training examples will be separated into the desired classes if α_i equals to zero, and will not be in the right classes if α_i is greater than *C*. Now, for the testing stage, a new data containing only the feature data, could be labeled by computing the sum

$$\mathbf{w}^\mathrm{T}\mathbf{z} + \mathrm{b} = \sum_{\mathrm{i}\in\mathrm{SV}} \alpha_\mathrm{i}\mathrm{y}_\mathrm{i}(\mathbf{x}_\mathbf{i}{}^\mathrm{T}\mathbf{z}) + \mathrm{b}, \tag{17}$$

which will classify $\mathbf{z}$ as the class $+1$ if the sum is positive and class -1 if the sum is negative.

For dealing with a more complex training data set, a non-linearity decision boundary is needed. Usually, a proper transformation to a higher dimensional feature space, where linear operation is equivalent to the non-linear operation in the input space, is desired. By transformation, the classifying task could be completed by applying the linear form of SVM.

B. Classification for Multiple Classes

In most recognition applications, multi-class classification tasks are required. We could combine several binary SVMs to accomplish a multi-class classification. There are two common strategies for solving a n-class classification problem [15] [20].

- By using the pair-wise strategy, $n(n-1)/2$ SVM classifiers are trained and each SVM separated a pair. This strategy requires a selecting mechanics to obtain final classification result. In the simplest case, majority voting method is adopted for the selecting mechanics. Other methods like the bottom-up tree structure and the top-down tree structure are also being proposed [21].
- By using the one-against-all strategy, n SVM classifiers are trained and each of the classifiers separates a class form the remaining in the training set. Notice in some cases, the one-against-all approach might result in ambiguous classification. Thus, we add a selection rule to avoid possible confusions. [21]

In the purposed system, we adopted the one-against-all strategies because less SVM classifiers are needed. Selection rules are designed as follow. When face features are identified as two or more classes or are not identified as any of the trained classes, we considered the identities of these faces as someone outside the trained database.

IV. TWO-CHANNEL FACE RECOGNITION

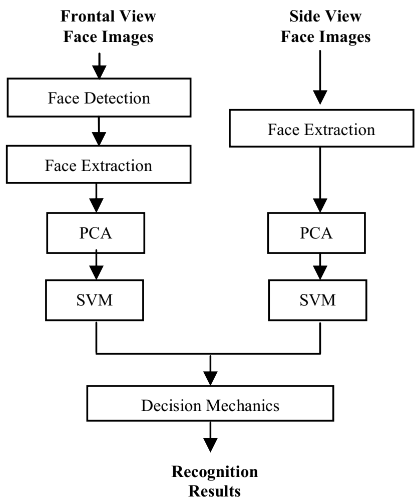

Figure 2. Flow chart of the proposed two channel face recognition systems.

The flowchart of the proposed two-channel face recognition system is shown in Fig. 2. Two still images, frontal faces and side faces, are required. Frontal faces are chosen because frontal view is the most distinguishable part of human faces among people. Side faces, on the other hand, are chosen because side views contain the information of depth and contours of human faces.

When a new pair of input images is entered, each channel performs its own recognition. Results from the two channels are fused by a decision mechanics, Bayesian interference, for a final result. In fact, a training process of the decision mechanics is necessary. Within each channel, a holistic 2D analysis on face images is performed, which will be described in the following.

A. Face Recognition

Before the face recognition stage, Adaboost based face detection is applied to frontal face images. A frontal face is extracted and its size is normalized. For a side face image, background subtraction is performed to eliminate miss recognition. Moreover, only lower part of the side faces are considered for face recognition.

The two input images are sent into their corresponding channels. A holistic 2D face analysis is performed within each channel. As mentioned in Section 2, PCA approaches are preferred in this case where the face database only contains limited amount of identities. Typically, minimum distance classifiers are employed for recognition in PCA approaches [14] [22]. This requires a large database and is very sensitive to illumination variations. In addition, misclassification often occurs when two identities look alike. Thus, SVM classifiers are applied instead of minimum distance classifiers in our system. For each face image, the components in the face space determined by PCA are defined as the features for the SVM classifiers. In other words, PCA are simply applied to reduce the dimension or to determine the features of each original face data and SVM are applied for classification.

For the SVM classification, the one-against-all strategy is preferred in the proposed system because less SVM classifiers are needed. A selection technique is designed to avoid an ambiguous result. For face features that are not identified as any classes or identified as more than one classes, they are considered as being someone outside the database.

B. Bayesian Fusion

There are many strategies to combine classification results. These strategies could be roughly divided into two categories, heuristics approaches and probability theory based approaches. Heuristics approaches include voting strategies, rank-based strategies and score-based strategies. [10] They are simple but less reliable. Thus, Bayesian theory based fusion proposed in [10] is adopted as the decision mechanics in our two-view face recognition system. Bayesian fusion considers individual recognition results of the two channels of each person, which makes it more reliable. In the following, we will state the algorithm and training process of Bayesian fusion.

Let $\boldsymbol{F}_j, j = 1, 2, \ldots, N$, denote the event that an input face image of person j has been shown to the channel with N being the number of classes. Let $\boldsymbol{H}_k$ denote the event that the result of the input image being identified as person k. The basic formula of Bayesian theorem is

$$P(\boldsymbol{F}_j|\boldsymbol{H}_k) = \frac{P(\boldsymbol{F}_j)P(\boldsymbol{H}_k|\boldsymbol{F}_j)}{\sum_{l=1}^{N} P(\boldsymbol{F}_l)P(\boldsymbol{H}_k|\boldsymbol{F}_l)}. \qquad (18)$$

Since the recognition processes of the two channels are independent, we could extend (18) to a more complex format. That is,

$$P(\boldsymbol{F}_j|\boldsymbol{H}_k{}^1 \cap \boldsymbol{H}_k{}^2) = \frac{P(\boldsymbol{F}_j)(P(\boldsymbol{H}_k{}^1|\boldsymbol{F}_j)P(\boldsymbol{H}_k{}^2|\boldsymbol{F}_j))}{\sum_{l=1}^{N} P(\boldsymbol{F}_l)(P(\boldsymbol{H}_k{}^1|\boldsymbol{F}_l)P(\boldsymbol{H}_k{}^2|\boldsymbol{F}_l))}. \qquad (19)$$

where $\boldsymbol{H}_k{}^i$ denotes the recognized identity from channel i, $i = 1,2$. Here, it is reasonable to assume that the prior probability $P(\boldsymbol{F}_j)$ is the same for each person j. That is, we set $P(\boldsymbol{F}_j) = \frac{1}{N}$ for every j.

To obtain $P(\boldsymbol{H}_k{}^i|\boldsymbol{F}_j)$ from each channel, a training process is required where training data must be large enough. For each channel a possibility set, $P(\boldsymbol{H}_k{}^i|\boldsymbol{F}_j)$, is determined by the possibility of each recognized identity $\boldsymbol{H}_k{}^i$ being the recognition result from an input image with the ground truth $\boldsymbol{F}_j$ in channel i. Once the possibility set is completed, we could calculate the joint decision, $P(\boldsymbol{F}_j|\boldsymbol{H}_k{}^1 \cap \boldsymbol{H}_k{}^2)$, of each $\boldsymbol{F}_j$. The $\boldsymbol{F}_j$ with the highest joint probability is elected as the final result.

V. EXPERIMENTAL RESULTS

Our experiments consist of two parts. The first part focuses on the front view channel. We compare our single channel face recognition with traditional PCA approaches using distance measurements. In the second part, we demonstrate the performance of our proposed two-channel system where its performance is compared with the two individual channels.

Figure 3. The camera setting for data collection: (a) Side view (b) Front view of the first data set by one webcam. (c) Front and side view of the second data set with two webcam.

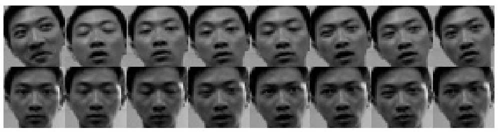

Figure 4. Examples of frontal view face data illuminations variations.

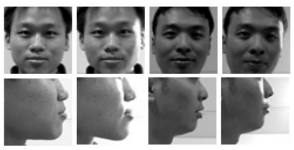

Figure 5. Examples of the face database with different face expressions.

Due to the lack of large amount of images with side view faces in the existing face databases, we must build a new face database containing both images with side view faces and frontal view faces for verifying our proposed two-channel face recognition. To minimize the storage size, we save only the extracted face images in the database. For the frontal views, face detection and face extraction, described in Section 4, are presented before setting up the databases. Face detection in this experiment is done by using the adaboost based detection module in OpenCV. For the side view channel, face extraction are completed before saving images into the database. The resolution of all face images in the database are 100×100. Examples of the frontal views and the side views and are shown in Fig. 5. Notice that the database contains several images with different illumination conditions and facial expressions in order to be more practical in real-world applications. Examples of frontal views with these changes are showing Fig. 4.

There are two sets of database. In the first set, facial images are captured one view at a time using only one webcam as shown in Fig 3 (a) (b). Two webcams are used for the second set as shown in Fig 3 (c). For each set, we collected three groups of face data for our experiments. The first group is the database of the face recognition of both channels, which is used to train the combined PCA and SVM. It consists of 10 persons with 2 illumination conditions and minor variation on pose and facial expression of each person. The second group is collected for training the decision mechanics. It tests the individual performance of each channel. This group consists of several lighting conditions and a wider range of variations on pose and facial expression for each person. Each of the two groups contains 400 images, 20 front view images and 20 side view images per person. The third data group is for testing purpose and it contains the 10 persons with wider illumination variations and facial expression changes. This testing group contains 350 pairs of face images of which 300 pairs are from the 10 persons.

A. Single-Channel Face Recognition

TABLE I. COMPARISIONS OF RECOGNITION RATE OF DIFFERENT PCA BASED APPROACHES. (13 EIGENFACES ARE SELECTED FOR THE SUBSPACE)

Face Recognition Scheme	PCA+ SVM	PCA+Weighted Manhattan distance	PCA+Weighted SSE distance
TPR Using the first data set	0.973	0.91	0.9
TPR Using the second data set	0.93	0.41	0.3933

In this experiment, we compare the performance of face recognition using a combination of PCA and SVM to that of the traditional PCA based approaches. We adopt two distance measure methods, weighted Manhattan distance and weighted SSE distance, from [22], to be incorporated with the traditional PCA based approaches. We use only the images with frontal views since most face recognition schemes are performed on this view. The first and second data set are used for the training stage and testing stage, respectively. We select 13 major eigenfaces in the feature space for face recognition. Table I shows the recognition results of PCA with SVM, PCA with weighted Manhattan distance, and PCA with weighted SSE distance. In this experiment, we have shown that the combination of PCA and SVM has a higher recognition rate than PCA approaches with distance measurement. Thus, we combine PCA and SVM in our proposed two-view face recognition system.

B. Two-Channel Face Recognition

In this part, we further compare the performance of the proposed two-view face recognition to those of the two individual single-channel face recognition. To verify the proposed system, a training stage for both face recognition of the two channels and the decision mechanics are needed. We used the first data set to train the combination of PCA and

TABLE II. COMPARISONS OF RECOGNITION RATES BETWEEN THE PROPOSED SYSTEM AND THE TWO INDIVIDUAL CHANNELS USING TWO DATA SETS

Face Recognition Scheme	Frontal-view Channel		Side-view Channel		Two-view System
TPR Using the first data set	Eigenface: 13	0.973333	Eigenface: 12	0.75	0.98
	Eigenface: 5	0.8567	Eigenface: 12	0.75	0.97
TPR Using the second data set	Eigenface: 11	0.93	Eigenface: 10	0.70	0.94
	Eigenface: 4	0.803333	Eigenface: 10	0.70	0.91
	Eigenface: 3	0.613333	Eigenface: 10	0.70	0.8333

TABLE III. ANALYSIS OF PERFORMANCE OF THE BAYESIAN FUSION IN THE FIRST DATA SET. (A TOTAL OF 350 PAIRS OF IMAGES FOR TESTING)

Individual Results from		Result of Bayesian Fusion	
Frontal	Side	Correct	Wrong
Correct	Correct	0.622857	0
Correct	Wrong	0.217143	0.017143
Wrong	Correct	0.125714	0
Wrong	Wrong	0.002857	0.014286

SVM. Only 10 pairs of each face identity are used. The second set is used to obtain the performance of the two channels individually and is also used for the training of the decision mechanics. Finally, we use the third data set to test the purposed system.

TABLE IV. ANALYSIS OF PERFORMANCE OF THE BAYESIAN FUSION IN THE SECOND DATA SET. (A TOTAL OF 350 PAIRS OF IMAGES FOR TESTING)

Individual Results from		Result of Bayesian Fusion	
Frontal	Side	Correct	Wrong
Correct	Correct	0.466667	0
Correct	Wrong	0.113333	0.033333
Wrong	Correct	0.213333	0.02
Wrong	Wrong	0.04	0.113333

Table II shows the comparisons of the recognition rates of the proposed system, frontal view channel and side view channel. In this part, we reduce the number of eigenfaces of the frontal view channel to see how the system reacts. It is shown in Table II that reducing the number of eigenfaces does little effect on the performance of the proposed system, even though the recognition rate of the frontal view channel became much lower. Moreover, we reduced the processing time and memory requirement in the face recognition stage. Table III and Table IV show the performance of the Bayesian theory based decision mechanics. Please notice that it is less likely for both channels to obtain a false recognition result at the same time. The fusion somehow corrects the channel with misclassification result. Also, there are several cases where both channel obtain a false recognition result but Bayesian fusion corrects it. This happens a lot while using the second data set. One person in the second set is easily being confused with a particular identity in the frontal channel and with another identity in second channel. While both confused results occurred on this person, the probability based fusion by (19), corrected the identity. In this experiment, it appears that the two channels somehow acting as a compensator to each other. Thus, the Bayesian fusion of the two-channel based face recognition really improves the recognition accuracy.

VI. CONCLUSION

In this paper, we proposed a two-view face recognition system to improve the recognition rate without involving with 3D analysis. The system fuses two individual recognition results from two channels with two input images with different viewing angles from a person based on Bayesian theory. Within each channel, a combined PCA and SVM are applied for face recognition. Our experimental results showed that this two-view face recognition system has achieved a higher recognition rate compared to traditional 2D single-view systems.

References

[1] W. Zhao, R. Chellappa, P. J. Phillips, and A. Rosenfeld, "Face recognition: a literature survey," *ACM Computing Surveys,* Vol. 35, No. 4, pp. 399-458, December 2003.

[2] B. Xie, V. Ramesh, Y. Zhu, and T. Boult, "On channel reliability measure training for multi-camera face recognition," *in proceedings of IEEE Workshop on Applications of Computer Vision,* February 2007.

[3] R. Zhang, P.-S. Tasi, J. E. Cryer, and M. Shah, "Shape from shading: A Survey," *IEEE Transactions on Pattern Analysis and Machine Intellignece,* Vol. 21, No. 8, August 1999.

[4] B. Achermann, X. Jiang, and H. Bunke, "Face recognition using range images," *in proceedings of International Conference on Virtual Systems and MultiMidia,* pp. 129-136, 1997.

[5] B. Xie, T. Boult, V. Ramesh, and Y. Zhu, "Multi-camera face recognition by reliability-based selection," *in proceedings of IEEE international Conference on Computational Intelligence for Homeland Security and Personal Safety, Alexandria, VA, USA*, October 2006.

[6] Y.n Lia, S. Gongb, J. Sherrahc, and H. Liddell, "Support Vector Machine based multi-view face detection and recognition," *Elsevier B.V., Image and Vision Computing,* Vol. 22, pp. 413-427, May 2004.

[7] P. J. Phillips, P. J. Flynn, T. Scruggs, K. W. Bowyer, J.Chang, K. Hoffman, J. Marques, J. Min, and W. Worek, "Overview of the face recognition grand challenge," *in proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition,* Vol.1, pp. 947-954, June 2005.

[8] G. Y. Zhang and J. Zhang, "Face recognition using multi-scale PCA and Support Vector Machine," *in Proceedings of 7th IEEE World Congress on Intelligent Control and Automation,* June 2008.

[9] V. Popovici and J.-P. Thiran "Face detection using an SVM trained in eigenfaces space," *in Proceedings of 4th Conference on Audio- and Video- biomatirc Person Authentication*, 2003.

[10] B. Achermann and H. Bunke, "Combination of classifiers on the decision level for face recognition," *University of Bern, IAM-96-002*, 1996.

[11] P. Belhumeur, J. Hespanha, and D. Kriegman, "Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* Vol. 19, No. 7, July 1997.

[12] A. M. Martinez and A. C. Kak, "PCA versus LDA," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* Vol. 23, No. 2, February 2001.

[13] K. Pearson, "On lines and planes of closest fit to systems of points in space," *The London Edinburgh and Dublin Philosophical Magazine and Journal of Science,* Sixth Series 2, pp. 559-572, 1901.

[14] M. Turk and A. Pentland, "Eigenfaces for recognition," *Journal of Cognitive Neuroscience*, Vol. *3*, No. 1, November 1991.

[15] W. Zhao and R. Chellappa, Face processing- advance modoling and methods, *Elsevier*, December 2005.

[16] C. Burgesa, "Tutorial on support vector machines for pattern recognition," *Springer, Data Mining and Knowledge Discovery* Vol. 2, pp.121-167, 1998.

[17] C. W. Hsu, C. C. Chang, and C. J. Lin, "A practical guide to support vector classification," *Department of Computer Science and Information Engineering, National Taiwan University, Taipei*, 2003.

[18] V. Vapnik, The Nature of Statistical Learning Theory, *Springer-Verlag, New York*, 1995.

[19] V. Vapnik, Estimation of Dependences Based on Empirical Data, *Nauka, Moscow*, 1979.

[20] P. Foggia, C. Sansone, F. Tortorella, and M. Vento, "Multiclassification: reject criteria for the Bayesian combiner," *Elsevier B.V. Pattern Recognition,* Vol. 32, pp. 1435-1477, August 1999.

[21] G. Guo, S. Z. Li, and K. Chan, "Face recogntion by support vector machines," *in Proceeding IEEE International Conference on Automatic Face and Gesture Recognition,* March 2000.

[22] V. Perlibakas, "Distance measures for PCA-based face recognition," *Elsevier B.V. Pattern Recognition Letters,* Vol. 25, pp. 711-724, April 2004.