A Study on Exact and Abstract Measurement
Face Classification Using Linear-Based Algorithms

September 2014

Chua Min Zhi
Dedication

To my family.
Abstract

Face classification is one of the subbranches of the image classification in computer vision. It includes a wide range of decision-theoretic approaches of machine learning to categorize face images into some predefined classes by analyzing the numerical properties of various face features. Depending on the predefined classes for the faces to be classified into, face classification can be categorized into 2 categories: (i) exact measurement face classification, and (ii) abstract measurement face classification. The exact measurement face classification has a crisp and clear definition of face classes, such as gender and identity, whereas the abstract measurement face classification has a vague definition of face classes, such as attractiveness. Among the different face classification approaches, linear-based approaches are known for their high accuracy performance, short processing time, and simplistic nature. This study proposes linear-based approaches in both exact measurement and abstract measurement face classification that are able to achieve high accuracy, and tolerable with face images in some reasonable variations in poses, lightings and face expression without compromising the speed of the algorithms.

Chapter 2 first presents an overview of the linear-based algorithms. Then, it describes Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA) and Fisher Linear Discriminant Analysis (FLD) in detail.

For the exact measurement face classification, this work deals with the face recognition, where the face classes are either known faces (users) and unknown faces (intruders) for face authentication or users' identities for face identification. Chapter 3 addresses the issues of the conventional FLD method and proposes a hybrid FLD methodology. The hybrid methodology combines the benefits of both global approaches that exploit whole face region for face features and local approaches that utilize local features, such as eyes, nose and mouth. In the proposed methodology, face images are divided into five components i.e., Fisher-face, Fisher-eyes, Fisher-eyesnose, Fisher-nosemouth and Fisher-mouth, to overcome the limitation of the conventional method in handling face images of an intruder. Results suggest that the proposed methodology is able to achieve 99.2% and 98.6% of accuracy in face authentication for the AT&T and MZ datasets, respectively. This methodology also yielded 97.5% and 96.3% of accuracy for face identification in the AT&T and MZ databases, respectively. The results of the proposed methodology are better than other conventional methods and other hybrid methods for both face authentication and face identification.

For the abstract measurement face classification, this work deals with the face classification based on personal attractive preference, where faces are categorized into attractive, common and non-attractive classes depending on how appealing the faces are, for each participant. Chapter 4 addresses the issues of the conventional PCA-based approaches for face classification based on attractiveness, and proposes a new methodology that uses class-specific eigenfaces and image reconstruction method to classify both Asian female and male faces based on personal attractive preference. The proposed methodology handles each class separately to obtain eigenfaces, which differs from the conventional PCA-based methods that trained face images in all the classes as one huge training set. Thus, the new methodology is able to produce eigenfaces that contain more specific face information for each class when compared to those produced by the conventional methods. Then, classification is performed by measuring the similarity between the test image and the reconstructed images using the class-specific eigenfaces. From the experiments, while the accuracy results vary depending on the participants’ personal attractive preference, the proposed method outperforms conventional methods for all the participants, with a confidence level of 95% according to the Wilcoxon signed-rank test. In the 3-class classification, the proposed method achieves improvement in average accuracy ranges from 7.7% to 15.1% and 2.9% to 17.4% for the
female and male datasets, respectively. Also, in the case of the 2-class classification, the improvement in average accuracy ranges from 6.3% to 17.7% and 5.2% to 16.6% for the female and male datasets, respectively. The proposed methodology not only achieves significant improvement in accuracy, its computational order still remains the same when compared with other conventional methods.

Finally, Chapter 5 summarizes this research, presents conclusions, and suggests future researches.
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.6 Conclusions</td>
<td>61</td>
</tr>
<tr>
<td>5 Conclusions and Future Works</td>
<td>63</td>
</tr>
<tr>
<td>5.1 Conclusions</td>
<td>64</td>
</tr>
<tr>
<td>5.2 Future Works</td>
<td>65</td>
</tr>
<tr>
<td>Bibliography</td>
<td>71</td>
</tr>
<tr>
<td>Publications and Awards</td>
<td>73</td>
</tr>
<tr>
<td>Acknowledgement</td>
<td>75</td>
</tr>
</tbody>
</table>
List of Figures

3.1 Example of false acceptance. ............................................. 13
3.2 Face component extraction for multimodal face recognition. ................. 14
3.3 Sub-holistic PCA. ................................................................... 14
3.4 Modular subspace LDA. .......................................................... 14
3.5 Example of how the proposed method can correct the false acceptance. ...... 16
3.6 Face component extracted for HFLD. ........................................ 17
3.7 Overview of the proposed methodology ...................................... 18
3.8 Face component extraction procedure (Preprocessing Stage) ................. 18
3.9 Preprocessing stage ............................................................... 20
3.10 Overview of the HFLD training stage ....................................... 21
3.11 Overview of the HFLD testing stage ....................................... 22
3.12 Algorithm flow of the HFLD training stage ............................... 23
3.13 Algorithm flow of the HFLD testing stage ............................... 24
3.14 Montage of face image examples from AT&T dataset ...................... 26
3.15 FAR and FRR versus different threshold values ............................ 27
3.16 ORR for proposed HFLD and other algorithms vs. different number of images per user in training training set ................................................. 30
3.17 IR for proposed HFLD and other algorithms vs. different number of images per user in training training set ................................................. 32
3.18 Montage of face image examples from MZ dataset ........................ 33
4.1 Conventional method vs. proposed method during training stage ............ 39
4.2 Conventional method vs. proposed method during testing stage ............. 40
4.3 Examples of eigenfaces for female dataset .................................... 41
4.4 Examples of eigenfaces for male dataset ..................................... 42
4.5 Block diagram of the conventional PCA-based face attractiveness classifier . 43
4.6 Block diagram of the proposed method ...................................... 45
4.7 Examples of reconstructed images for female dataset ....................... 46
4.8 Examples of reconstructed images for male dataset ........................ 47
4.9 Examples of face images in the female dataset ............................... 48
4.10 Examples of face images in the male dataset ................................ 48
4.11 Results for 3-class female classification ..................................... 52
4.12 Results for 3-class male classification ...................................... 52
4.13 Confusion matrix for female dataset ....................................... 54
4.14 Confusion matrix for male dataset ........................................... 55
4.15 Results for 2-class female classification ..................................... 57
4.16 Results for 2-class male classification ...................................... 57
4.17 Results for hybrid version of the proposed method ........................ 59
# List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>Dimensions after preprocessing</td>
<td>19</td>
</tr>
<tr>
<td>3.2</td>
<td>Optimum value of $\alpha$</td>
<td>27</td>
</tr>
<tr>
<td>3.3</td>
<td>Comparison result of different classifier</td>
<td>28</td>
</tr>
<tr>
<td>3.4</td>
<td>Comparison of results for different voting scores</td>
<td>28</td>
</tr>
<tr>
<td>3.5</td>
<td>Comparison results based on $k$-nnr Classifier</td>
<td>29</td>
</tr>
<tr>
<td>3.6</td>
<td>Comparisons of recognition rates, error rates and computation time for Experiment II</td>
<td>29</td>
</tr>
<tr>
<td>3.7</td>
<td>Comparisons of recognition rates, error rates and computation time for Experiment III</td>
<td>31</td>
</tr>
<tr>
<td>3.8</td>
<td>Comparison results for Experiment IV</td>
<td>31</td>
</tr>
<tr>
<td>3.9</td>
<td>Comparison results for Experiment V</td>
<td>32</td>
</tr>
<tr>
<td>3.10</td>
<td>Comparison of results for Experiment VI and Experiment VII</td>
<td>34</td>
</tr>
<tr>
<td>4.1</td>
<td>Algorithm List</td>
<td>49</td>
</tr>
<tr>
<td>4.2</td>
<td>Averaged improvement in accuracy ($\bar{A}_q$) of the proposed method for 3-class classification (in percentage)</td>
<td>53</td>
</tr>
<tr>
<td>4.3</td>
<td>Averaged improvement in accuracy ($\bar{A}_q$) of the proposed method for 2-class classification (in percentage)</td>
<td>56</td>
</tr>
<tr>
<td>4.4</td>
<td>Experiment List</td>
<td>59</td>
</tr>
<tr>
<td>4.5</td>
<td>Averaged improvement in accuracy ($\bar{A}_q$) of the proposed method (Experiment 1) with different variations of the hybrid version proposed method (in percentage)</td>
<td>60</td>
</tr>
<tr>
<td>4.6</td>
<td>Computational Time</td>
<td>61</td>
</tr>
</tbody>
</table>
Chapter 1

Introduction

This chapter presents a brief introduction to exact and abstract measurement face classification. Then, it describes the linear-based approaches and other approaches. Followed by the descriptions on the challenges and goals of this research. Finally, it outlines the content of this work.
1.1 Image-based Classification

Image classification [1, 2, 3, 4, 5, 6] techniques in computer vision include a wide range of decision-theoretic approaches of machine learning. It can be defined as a process to categorize the spectral data of a single image or a series of images into their corresponding classes. Classification is performed by analyzing the numerical properties of various image features. There are two main branches of classification approaches, i.e., supervised classification and unsupervised classification. In the supervised classification approach, the classes are predefined based on given criteria. On the contrary, the unsupervised classification approach first examines pixels on the test image, and then, classification is accomplished by the natural groupings based on the pixel values within the area of interest of the test image. In this thesis, we only focus on supervised image classification.

A classification approach typically employs two main stages of processing, namely, (i) training stage and (ii) testing stage. In the training stage, useful characteristic properties are extracted and descriptions of the training images are created. In the subsequent testing stage, the extracted features are utilized to classify some given test images.

There are many important applications for the numerous image classification techniques in different areas that are directly affecting our daily lives. For examples, face recognition security system [7] for security purposes, land cover classification using satellite imagery [8] for city planning and development, medical image classification [9] for early cancer or other disease detection, and hyperspectral remote sensing image classification [10] for nature observation on earth.

Depending on the desired criteria for the classification, the image classification could be subcategorized into two groups, i.e, (i) the exact measurement classification and (ii) the abstract measurement classification. The terms ‘exact’ and ‘abstract’ measurements indicate the ambiguity of the predefined classes. In the exact measurement classification, the classes are well defined and unambiguous. Some examples of these type of classifications are face recognition [7], gender classification [11], image genre classification [12] and facial expression classification [13]. On the other hand, in the abstract measurement classification, the classes are more ambiguous, for example, aesthetic visual quality assessment of paintings [14], photo assessment [15], and emotion portrayed from a natural image [16], where the classes for classification are based on feelings or emotions that are vague and therefore harder to measure.

1.2 Face Classification

Face classification is a popular subbranch of the image classification systems in which the region of interests are faces within the image. There are numerous ways to classify a face, for example, face recognition [7], face classification based on gender group [11], age [17], and ethnicity [18]. The decisions of these type of face classifications are either human faces or background object, a known face or an unknown face, male or female, and etc. There are no ambiguity between the classes, therefore, in this study, the face classification with these unambiguous classes will be referred to as exact measurement face classification.

Contrast to the exact measurement face classification that gained the attention from researchers for the past years, abstract classes of classification have just recently became the new challenge for researchers in this area. The relatively new research in abstract measurement for face classification is the face classification based on attractiveness[19, 20, 21, 22, 23]. In this area of research, face images are classified into different classes based on how appealing the faces are. This measurement of the attractiveness of the face classes are ambiguous and abstract. Therefore, in this
research, the face classification with these vague classes will be referred to as *abstract measurement* face classification.

### 1.3 Linear-Based Approaches and Others

Numerous approaches for face classification was developed during past years. The different approaches vary from the geometric local landmark method that sometimes requires manual pin point frontal view of facial parameters from a face image [24, 25, 26] to the more computational expensive neural network algorithms [27, 28, 29].

Among the existing techniques, PCA approach is one of the popular approach. Sirovich et al. [30] are the first to introduce that the face images can be representation in lower dimensionality using the concept from Principal Component Analysis (PCA). Since then, the concept of eigenpictures has became one of the most popular methodologies in the area of face-based computer vision such as face representation, face detection and face classification. They demonstrated that face images can be reconstructed using linear summation of the picture form of eigenpictures (the eigenvectors) and their respective coefficients.

Turk and Pentland [31] later adapted the idea from Kirch and Sirovich and successfully demonstrated an automatic face recognition system using the eigenpictures (or known later as eigenfaces) for face detection and face identification [32]. In eigenfaces, each face from the dataset is represented as vector of weights that are obtained from projection of that image onto the eigenface plane. During the recognition stage, the input image is projected to the eigenface plane and the weights are compared with those obtained from the dataset images using simple Euclidean distance. Locating the nearest weights will then identify the input image. This is the most basic form of linear-based approach for face classification.

This basic form is extended to the Bayesian approach [33, 34] that requires the estimation of probability distributions in a high-dimensional space. And later to a nonlinear form of PCA that is also known as the kernel PCA proposed by Schölkopf et al. [35]. There are also other approaches that involve PCA to reduce the dimensions of the extracted features. As an example, Liu et al. [36] presented the Evolution Pursuit (EP) method that first projects the optimal basis face features extracted onto a lower-dimensional PCA space, and then implements genetic algorithms (GA) for face classification.

Another successful linear based approach is the Linear Discriminant Analysis (LDA). While PCA trains the optimized distribution for the entire face features, LDA trains the optimized distribution based on classes. It distinguishes within-class and between-class variations. This concept is first introduced by Belhumeur et al. [37] introduced the conventional LDA. Small sample size (SSS) [38] problem occurs when the within-class scatter matrix becomes singular which is often the case for high dimensional data such as face images. Belhumeur et al. [37] also proposed Fisher Linear Discriminant Analysis (FLD) and that involves both LDA and PCA to overcome the issue. FLD avoids the complication of a singular within-class scatter matrix by projecting the image set to a lower dimensional space using PCA.

There are many alternative solutions proposed to solve the SSS issue. Chen et al. [39] proposed Chen-LDA (CLDA) to overcome the singularity problem. They use the discrimininative information from the null space of the within-class scatter matrix instead whenever it is singular. Zhao et al. [40, 41] proposed another methodology that first projects the face dataset onto the PCA projection to reduce its dimensions, and then followed by LDA projection. The PCA step helps to remove null spaces from the within-class scatter matrix and the between-class scatter matrix to avoid SSS problem.
Yang et al. \cite{yang2007} and Yu et al. \cite{yu2008} presented a new direct-LDA (D-LDA) algorithm that maximizes the LDA criterion directly by diagonalizing between-class scatter matrix first instead of within-class scatter matrix when solving the generalized eigenproblem. These approaches avoid losing the most discriminant information by eliminating the requirement of the feature extraction or dimensionality reduction steps such as the separate PCA step that removes null spaces from within-class scatter matrix. Juwei et al. \cite{juwei2009} proposed an extended idea from D-LDA called the Direct Fractional-Step LDA (DF-LDA) algorithm to solve the extreme case when there is only a single sample in a class for LDA algorithm. This is accomplished by doing the D-LDA procedure followed by fractional step to further reduce the dimensionality of the training set.

There are various form of LDA-based methods available. Zhou and Chellappa \cite{zhou2007} proposed Multiple-Exemplar Discriminant Analysis (MEDA) that uses multiple face average images to represent a single class instead of a single face average image per class in the conventional LDA approaches.

Other than the PCA-based and LDA-based approaches, there are also other models that have been proposed, such as the Independent Component Analysis (ICA) Bartlett et al. \cite{bartlett2006} that introduced statistically independent basis to represents images, the Hidden Markov Models (HMM) \cite{hmm1,hmm2}, biologically inspired Local Feature Approach models (LFA) \cite{lfa}, flexible appearance models \cite{flexibleappearance}, Elastic Bunch Graph Matching Models (EBGM) \cite{ebgm} and neural network approaches \cite{neuralnetwork1,neuralnetwork2,neuralnetwork3}.

Among the different approaches mentioned above, the linear-based algorithms, i.e., linear PCA-based and linear LDA-based approaches appeared to be very successful methods in complex task that involved high dimension data such as face images \cite{faceimages1,faceimages2}, in spite of their simplistic nature. These methods are also the basis of many success methods for face classification applications.

1.4 Challenges and Goals

There are some characteristics that are important for face classification applications. A suitable approach should be able to do the following:

- achieve high classification accuracy.
- complete the complicated task within a reasonable time.
- tolerate face images with some reasonable variations in poses, lightings, face expression and etc.

Linear-based approaches have shown tremendous success in fulfilling all the basic requirements mentioned previously. Due to the impressive simplicity yet powerful classification capability, this study will be focusing on the linear PCA-based and linear LDA-based algorithms. The main objective of this thesis is to propose methodologies to improve the performance of these linear-based algorithms without compromising the speed of the algorithms. However, the requirements and importances of these characteristics differ depending on the application of the face classification. As mentioned in section 1.2, face classification can be divided into the exact and abstract measurement face classification. For the exact measurement face classification, this study will focus on face recognition for security. And for the abstract measurement face classification, this study will focus on the face classification based on attractiveness.

The two most important characteristics for the approach used in the face recognition security system is that it should be able to achieve high accuracy with short processing time. According to
[52], an accuracy rate of more than 90% is considered as a good rate for general applications, but for approaches used in security application, the accuracy should be much higher. The linear-based algorithms mentioned are able to achieve relatively high accuracy. However, as the reliability of the face recognition security system is directly proportional to the accuracy that the approach can achieve. Therefore, the main goal for this study is to propose a new methodology that can not only able to outperform the other popular linear-based algorithms, also to aim for accuracy rate closer to 100%. There are 2 major tasks for the face recognition security system, (i) to prevent intruders, and (ii) to correctly distinguish the identity of each individuals within the training set. An effective face recognition security system should be able to perform these 2 tasks using the same algorithm. Also, there are 3 ways to extract pixel intensity as input features from a face image, (i) global approach that takes the entire face as input, (ii) local approach that takes local features such as the eyes and nose as input, and (iii) hybrid approach that is the combination of both. To take the advantage of the global and local approaches, this study proposes a hybrid face recognition methodology that is able to achieve high accuracy in the 2 major tasks of face recognition security system. For a practical security system, other than testing the accuracy to ensure security, this study also tests the computational time to ensure that the new methodology is able to complete the task within a reasonable period of time, i.e., within seconds. The face images used for testing varies in lightings, poses, facial expressions and accessories to examine the tolerancy of the approach.

Face classification based on attractiveness is a relatively new research branch. There are a few approaches that have been proposed but there are still many areas that have not been explored. To the best of our knowledge, the approaches proposed are designed specifically and tested with only female faces. Also, individual opinions are neglected and the face images are classified based on the general perspective of the term attractiveness. The linear-based conventional approaches proposed by other researchers used the PCA algorithm. The conventional PCA-based approaches proposed for the face classification based on attractiveness are utilizing eigenfaces that is the output from training the entire dataset together. These general eigenfaces did not utilize the class information and were not specific to represent face images from each class separately. Therefore, the main goal for this study is to propose a class-specific eigenfaces methodology that creates unique and specific eigenfaces for each class separately. This methodology should able to produce a better representations of the face images from the training set when compares with the conventional approaches. The new methodology is expecting to be non gender specific, able to outperform the other conventional PCA-based approaches in terms of accuracy, and yet maintaining the sophisticated simplistic nature of a linear-based approach. The new methodology should take each individual’s personal attractive preference into consideration for profitable applications such as automatic match making system. Here, this study tests the new methodology with both female and male face images. The face images for testing varies in lightings, poses, and facial expressions to examine the tolerancy of the approach.

1.5 Outline

The central theme of this work is the exact measurement and abstract face classification using linear-based algorithms.

Chapter 2 describes the basic linear-based algorithms that are the main focus in this study, i.e., PCA, LDA, and FLD.

Chapter 3 presents a hybrid version of Fisher Linear Discriminant Analysis method (HFLD) for the exact measurement face classification system, i.e., face recognition system. Firstly, it
explains in detail about the mentioned drawback in FLD. Then, it describes the detailed overview of this methodology. To evaluate the proposed method, this chapter also presents experiments and results comparing the effectiveness of the proposed method with other algorithms.

**Chapter 4** describe a relatively new abstract measurement face classification approach, i.e., face classification based on attractiveness. Firstly, it introduces the concept of the face classification based on personal attractive preference. Then, it explains the issues that motivate us to propose the face classification based on personal attractive preference using class-specific eigenfaces reconstruction method, follows by a detailed overview of the proposed methodology. To evaluate the proposed method, this chapter also presents experiments and results comparing the effectiveness of the proposed method with other conventional approaches.

Finally, **Chapter 5** summarizes this research, presents conclusions, and suggests future researches.
Chapter 2

Overview of the Linear-Based Algorithms

This chapter presents a brief introduction for the linear-based algorithms. Then, it describes the PCA, LDA and FLD, which are the main focus of this study.
2.1 Introduction

Linear-based algorithms such as Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA) and Fisher Linear Discriminant Analysis (FLD) have shown tremendous success in various applications for image classification despite their simplicity. These approaches are very suitable for realistic application in image-based face classification because they fulfill the essential conditions as suitable algorithm for face classification, i.e., (1) they are able to achieve high classification accuracy, (2) they are not computationally expensive, and (3) they are robust towards face image variations. Many algorithms that are based on the enhancement or improvement on these basic linear-based algorithms have been proposed throughout the years. The following sections present overviews of the fundamental linear-based algorithms that are of interest in this study.

2.2 Principal Component Analysis

PCA is widely used in high dimensional data applications, and was first introduced into image processing applications such as face recognition by Turk and Pentland [31]. Usages of PCA include dimensionality reduction, image compression, image reconstruction, image recognition and image classification.

PCA is performed through the eigen decomposition of the covariance matrix of images. Suppose that we have \( n \) 2D gray scale images with \( J \) rows and \( K \) columns, and \( m = JK \) pixels in total. Let \( x_i \in [0,255] \) be a \( m \times 1 \) column vector consisting of the intensities of all the pixels converted from the \( i \)-th 2D gray scale image. Each new pixel position \( j_{\text{new}} \) of \( x_i \) is computed as follows

\[
j_{\text{new}} = J(k - 1) - j,
\]

for \( 1 \leq j_{\text{new}} \leq m \), \( 1 \leq j \leq J \), and \( 1 \leq k \leq K \), where \( j \) and \( k \) denote the row and column positions of a pixel in the 2D gray scale image. Let \( X = [x_1, \ldots, x_n] \) be the data matrix of the column vector images, \( \mu = \frac{1}{n} \sum_{i=1}^{n} x_i \in \mathbb{R}^m \) be the average of all the images, i.e., the average of the rows of \( X \). Let \( \psi_i = x_i - \mu \) be the deviation of \( i \)-th images from the mean image \( \mu \), and \( A = [\psi_1, \ldots, \psi_n] \). Then, the covariance matrix of the set of the images is given by

\[
C = \frac{1}{n} AA^T.
\]

The eigen decomposition of \( C \) is \( C = EDE^T \), where \( D \) is a diagonal matrix whose diagonal elements are the eigenvalues of \( C \) and \( E \) is an orthogonal matrix whose \( i \)-th column is the normalized eigenvector corresponding to the \( i \)-th diagonal element (eigenvalue) of \( D \).

Since \( C \) has dimension of \( m \times m \), the cost of the eigen decomposition usually requires a computational complexity of \( m^3 \) to obtain \( D \) and \( U \). Fortunately, in this case, since the dimension of \( A \) is \( m \times n \) where \( n \ll m \) in many situations, the covariance matrix has at most \( n \) nonzero eigenvalues. Notably, nonzero eigenvalues of \( C \) can be computed through the singular value decomposition (SVD) of \( A \), which is \( A = U \Lambda V^T \), where \( \Lambda \) is a \( m \times n \) diagonal matrix whose diagonal elements are the singular values of \( A \), \( U \) is a \( m \times m \) matrix whose columns have norm 1 and are orthogonal to each other, and \( V \) is a \( n \times n \) dimensional orthogonal matrix. Note that \( EDE^T = C = \frac{1}{n} AA^T = \frac{1}{n} U \Lambda V^T \Lambda U^T = \frac{1}{n} U \Lambda A^2 U^T = \frac{1}{n} U A^2 U^T \), meaning that diagonal elements of \( \Lambda^2 \) correspond to nonzero diagonal elements of \( D \) and the columns of \( U \) are equal to the columns of \( \Lambda \) whose corresponding eigenvalues are nonzero. The computational complexity of the SVD is in \( O(n^2m + n^3) \) [53] and is dramatically smaller than \( m^3 \) when \( n \ll m \). We call each column of \( U \) an eigenface.
In PCA, the eigenfaces \(U\) is chosen to maximize the determinant of the covariance matrix (or also known as the scatter matrix) of the trained images \(X\), i.e.,

\[
U = \arg\max_U |U^T C U|
\]

\[
= [u_1 \ u_2 \ \ldots \ u_n],
\]

where \(u_i\) for \(i = 1, 2, \ldots, n\) is the set of \(m\)-dimensional eigenfaces of \(C\) corresponding to the \(n\) largest eigenvalues.

Once we obtain \(U\), all the images in the dataset are expressed as linear combinations of columns of \(U\). Moreover, any linear combination of the images in the dataset, e.g. \(\mu\), is a linear combination of columns of \(U\). Given an image vector \(y\), the projected image \(y'\) into the space spanned by the columns of \(U\) is

\[
y' = \mu + U \Omega_i,
\]

where

\[
\Omega_i = U^T \psi_i.
\]

We call \(\Omega_i\) as the weight corresponding to \(x_i\).

### 2.3 Linear Discriminant Analysis

LDA [37] produces an optimal linear discriminant function \(f(x) = V^T x\) that maps the input into the classification space. It seeks the best directions that are the most effective for data discrimination. Suppose we have \(p\) number of classes, each class contains \(q\) number of gray scale images and each image contains \(m\) pixels. Let \(x_i \in [0, 255]\) be the vector consisting of the intensities of all the pixels in the \(i\)th image. Let \(X_j = [x_{1,j}, \ldots, x_{q,j}]\) be the data matrix for class \(j\), \(\mu_o\) be the average of all the images, and \(\mu_j\) be the average of images in class \(j\), i.e., the average of the rows of \(X_j\). The within-class scatter covariance matrix \(S_w\) and between-class scatter covariance matrices \(S_b\) are given by

\[
S_w = \frac{1}{q} \sum_{i=1}^{p} \sum_{j=1}^{q} (x_{i,j} - \mu_j)(x_{i,j} - \mu_j)^T,
\]

\[
S_b = \frac{1}{p} \sum_{j=1}^{p} (\mu_j - \mu_o)(\mu_j - \mu_o)^T,
\]

where

\[
\mu_o = \frac{1}{pq} \sum_{j=1}^{p} \sum_{i=1}^{q} x_{i,j},
\]

and

\[
\mu_j = \frac{1}{q} \sum_{i=1}^{q} x_{i,j}.
\]

Given that \(S_w\) is non singular, i.e., \(q > 1\), the optimum projection \(V\) is chosen to maximize the ratio of determinant of the \(S_b\) to \(S_w\) [37] by solving the generalized eigenvalue problem

\[
S_b W = \lambda S_w W.
\]
We call each column of \( V \) an eigenvector for LDA. In LDA, \( V \) is chosen as the matrix with orthonormal columns that maximizes the ratio of the determinant of the between-class scatter matrix to the determinant of the within-class scatter matrix of the trained images \( X \), i.e.,

\[
V = \arg \max_V \frac{|V^T S_b V|}{|V^T S_w V|}
\]

\( (2.11) \)

where \( v_i \) for \( i = 1, 2, \ldots, n \) is the set of \( m \)-dimensional eigenvectors of LDA corresponding to the \( n \) largest eigenvalues.

The main issue confronted by the LDA in the face recognition problem is that the within-class scatter covariance matrix \( S_w \) is always singular \[37\]. Therefore the direct usage of LDA in the problem is not possible. The following section describes a solution for the mentioned issue.

2.4 Fisher Linear Discriminant Analysis

In order to ensure a nonsingular within-class scatter covariance matrix, PCA is utilized first to reduce the dimensionality of the feature space \[37\] then the standard LDA defined by Eqn. 2.3 is applied. The fisherfaces \( W \) can be obtained by solving the generalized eigenvalue problem as shown:

\[
S_{bb} W = \lambda S_{ww} W
\]

\( (2.12) \)

where

\[
S_{bb} = U^T S_b U \quad \text{and} \quad S_{ww} = U^T S_w U
\]

\( (2.13) \)

where \( U \) is the eigenfaces defined from section 2.2.

We call each column of \( W \) an fisherface. In FLD, \( W \) is chosen as the matrix with orthonormal columns that maximizes the ratio of the determinant of the \( S_{bb} \) to the determinant of the \( S_{ww} \) of the trained images \( X \), i.e.,

\[
W = \arg \max_W \frac{|W^T U^T S_b U W|}{|W^T U^T S_w U W|}
\]

\( (2.14) \)

where \( w_i \) for \( i = 1, 2, \ldots, n \) is the set of \( m \)-dimensional fisherfaces corresponding to the \( n \) largest eigenvalues.

2.5 Conclusion

Fundamental overviews of the linear-based algorithms that will be further investigates in the exact measurement and abstract measurement face classification in later chapters were presented.
Chapter 3

Exact Measurement Face Classification

This chapter presents an overview for one of the exact measurement face classification, i.e., face recognition approach to challenge the performance of the linear-based algorithm that we propose. This chapter begins with an introduction. Then, it describes the advantages and drawback in the conventional FLD method and subsequently, proposes a hybrid FLD methodology. Experiments are carried out to verify the relative performance of the proposed method with respect to the other global approach PCA and LDA methods as well as other hybrid approach algorithms.
3.1 Introduction

The exact measurement face classification is defined as an approach that classifies faces into pre-defined classes with clear and unambiguous definition. Gender classification [11] that classifies faces into male and female classes, age classification [17] that classifies faces into different age range and face recognition system that recognizes and classify faces into individual identities are some of the examples in this category.

There are numerous exact measurement face classifications, however, this study challenges various linear-based algorithms in face classification based on identify of an individual, or more broadly known as face recognition because there are a lot of possible revenue generating applications in this branch of research. One of the most important application for face recognition is face recognition security system. Face recognition security system is categorized as the vision-based biometric security system. The system is able to do perform in long range and does not required corporation from the subject. It detects and identifies intruders from a range without alerting the intruders. This advantage encourages the development of various type of face recognition security system.

There are a few important characteristics that the approaches for this application should possessed, i.e., (i) high recognition rate, (ii) short processing time and (iii) robust towards situation, i.e., insensitive to posing and lightings of the image. Firstly, the approaches should have high recognition rate. According to [52], a recognition rate of more than 90% can be considered as a good rate. However, for security purposes, the ideal recognition rate should be near to 100%. Therefore, other security methods such as fingerprint recognition, user’s name and password, and etc will be paired with the visual type security system to compromise with the recognition rate if necessary and strengthen the security.

Secondly, the approach should be able to derived a decision in relatively short processing time, i.e., within seconds. Computationally expensive algorithm is not practical for face recognition security system because the user should be able to gain or denied access in a reasonable time. And thirdly, the approach should be robust towards the input image with various background. Robustness is the flexibility of the algorithm to be insensitive to shape, pose and light. This characteristic is crucial to obtain high recognition rate. Also, the characteristic is very important for practicality reason as the approach should be able to process face images that obtained from various background environment under different lightings and with image of the faces taken in some angle within certain range.

For a flexible face recognition system, the approach should be able to classify individuals’ faces accurately, but also able to recognize an unknown face. From here onwards, the process of accepting known faces from the training set and rejecting unknown faces will be referred to as face authentication; the process of classifying different individuals into their corresponding identity will be referred to as face identification; individuals that enrolled into the training stage will be referred to as users; and, individuals that did not undergo the training stage will be referred to as intruders.

Among the aforementioned basic linear-based approaches, Fisher Linear Discriminant Analysis (FLD) method that involved both linear-based PCA and LDA appears to be an attractive choice because it is simple, easy to implement, fast processing time, and robust. The approach has the reputation of achieving higher accuracy in recognizing faces comparing to PCA approach because it considers the class variation information within the training images. However, it does not perform as well when faces that did not been through the training stage are presented to it [41] because FLD assumes that the test image is from the training set and still try to identify it as one of the users. Figure. 3.1 illustrates an example of face-spaces of two users from the training set,
Figure 3.1 Example of false acceptance. Weight distributions of Fisher-face space for user A (denoted by ‘×’), user B (denoted by ‘O’) and intruder X (denoted by ‘Δ’).

i.e., user A and user B, and are bounded by Fisher-face vectors, \( W_{\text{face}1} \), \( W_{\text{face}2} \) and \( W_{\text{face}3} \). These regions specify the weight distributions of each trained image of the users. A test image of an intruder X is projected onto the Fisher-face space. The FLD classifier will tries to determine whether the face belongs to user A or user B. In this example, this intruder is misclassified as user A, which consequently results in a false acceptance case.

This drawback motivates us to propose a Hybrid Fisher Linear Discriminant (HFLD) analysis algorithm that is able to improve the accuracy of the algorithm when handling unknown face images while maintaining the high accuracy of classifying the users into their correct identity.

### 3.2 Background

The concept of a hybrid face recognition algorithm that fuses both global and local approaches was first introduced by Pentland et al. [54]. In the work, they proposed modular eigenspaces for face recognition that uses both global eigenfaces and local eigenfeatures such as eigenmouth and eigeneyes. Their experiments claimed that the feature-based concept is useful when there are variations in expressions and face decorations in the input images.

This concept is later adopted by Guangda et al. [55] in their multimodal face recognition method (MMPCA) for simulated portrait. They extracted images of pure face, brows and eyes, eyes, nose and mouth from a geometrically normalized face image and PCA is used for dimensional reduction. This algorithm is applied in simulated portrait recognition to recognize criminals. Figure 3.2 shows the facial parts of the multimodal face recognition method.

Penev et al. [49] developed a biologically inspired feature analysis method called Local Feature Approach (LFA). LFA is used to extract topographic local features from global PCA algorithm. The method searches for the best topographic set of sparsely distributed grids based on reconstruction error. This method is used in one of the commercial system Facelt.

Lanitis et al. [50] proposed a flexible appearance model based method for automatic face
Figure 3.2 Face component extraction for multimodal face recognition.

Figure 3.3 Sub-holistic PCA.

Figure 3.4 Modular subspace LDA.
recognition. In this method, both shape and gray-level information are modeled and used to identify a face. The shape model is trained on example images using PCA, where the variables are the coordinates of the shape model points. The shape variations due to inter-class variation are separated from those due to within-class variations for classification. Consequently, a global shape-free gray-level model can be constructed using PCA.

Huang et al. [56] proposed a component-based method to decompose a face into a set of facial components such as mouth and eyes that are interconnected by a flexible geometrical model. However, a major drawback of this system is that it requires large number of training images from different viewpoints under different lighting conditions. To overcome this issue, the 3D morphable face model by Blanz et al. [57] is applied to generate arbitrary synthetic images under varying pose and illumination. Further researches on this approach are done by [58] and [59].

Khan et al. [60] proposed a sub-holistic PCA (SH-PCA). SH-PCA takes not only the whole face image but also four sub-images from one fourth of the image from each corner. From the experimental data, SH-PCA yields improvement of 6% compared to traditional PCA. Figure 3.3 shows how to separate the four sub-images from the whole face image.

Price et al. [61] proposed a modular subspace for LDA. The framework employs parallel system of observers; each is trained on a specific region of the face. Each observer is a linear subspace classifier and the output of all the observers are combined using a simple sum-rule. Figure 3.4 shows the image separation of modular subspace for LDA.

3.3 Methodology

This section discusses the concept of the proposed method. Then, it presents the overview of the proposed method.

3.3.1 Concept of the Proposed Method

This face division methodology will convert the global approach of the FLD into a hybrid approach. The proposed HFLD is able to make a decision on whether the input face is from the training set or not as shown in Figure 3.5. By dividing the face image into subsections, if the input face image is from an intruder, it is unlikely that it will match with the same person for all subsections. After projecting the input image of intruder X into subsectioned Fisher spaces, the weights for each space are compared. The distance in the face space of the intruder X in Figure 3.5(a) is closer to user A by chance. Therefore, the matching score for the Fisher-face space is user A. However, in the Fisher-eyes space of the intruder X shown in Figure 3.5(b), it is nearer to user B, therefore, the matching score for the Fisher-eyes space is user B. Due of the inconsistency in the matching score, the test image will be classified as an intruder.
Figure 3.5 Example of how the proposed method can correct the false acceptance. (a) Weight distributions of Fisher-face space for user A (denoted by ‘×’), user B (denoted by ‘O’) and intruder X (denoted by ‘Δ’). (b) Weight distributions of Fisher-eyes space.
Section 3.2 discussed different ways to divide a face image. However, the aforementioned methods do not utilize face features in the most optimized manner. According to Shepherd et al. [62], hair, face outline, eyes, and mouth are the significant facial features for human beings to perceive and remember faces. Bruce et al. [63] researched on the importance of the nose. He observed that the nose is insignificant in facial recognition except for recognizing face profiles. Following this, the proposed HFLD will not treat the nose as a single input image, but will combine it with other features. It is known that the nose alone will make it quite impossible for humans to recognize each other. However, if a person has one of his or her other features such as the eyes or mouth along with the nose, as shown in Figure 3.6(d) and Figure 3.6(e), he or she is still recognizable. Since majority voting method is used to classify HFLD, an odd number of components are required to avoid a tie in the score.

It is given that a person’s appearance will change slightly on a daily basis. For example, hair, moustache and beard will grow or change styles, different accessories like sunglasses, nose rings, and bandanna can also be worn depending on the day’s outfit. Cropping the whole face of the input face image will eliminate the effect of changes on the hair style, hence, improving the robustness of the algorithm. Cropping the whole face into several components will reduce the effects of illumination, expression and decorations which will subsequently increase the recognition rate of the algorithm.

Consequently, we propose a novel methodology that divides the input face image into five subsections: whole face, eyes, eyes and nose, nose and mouth, and mouth as shown in Figure 3.6.

### 3.3.2 Proposed Methodology

The proposed methodology is divided into two stages: (i) training stage, and (ii) testing stage. Before each stage, face images undergo a preprocessing stage to divide the face images into 5 subsections. Figure 3.7 illustrates the overview of the proposed methodology.

#### Face Component Extraction

The preprocessing stage in the proposed methodology can be referred to as the face component extraction. Consider that a gray scale face image as a 2D \( x \times y \) matrix, where \( x \) and \( y \) denote the coordinates of each pixel. The first step in the preprocessing stage is to extract the face features into five subsections mentioned. The process begins with the eye region. After determining the iris center point for both the left eye, \((x_l, y_l)\), and the right eye, \((x_r, y_r)\) using a template matching method after [26], the distance between the two points are measured and denoted as \( I_d \). Through trials and experiments, the following anthropometric measures were used for a consistent and
reliable approximation of the eye region, denoted $R_{\text{eye}}$.

Top eye boundary:

$$x_{\text{TB}} = \left( \frac{x_l + x_r}{2} \right) - \frac{1}{4} \times I_d \, , \quad (3.1)$$

Bottom eye boundary:

$$x_{\text{BB}} = \left( \frac{x_l + x_r}{2} \right) + \frac{1}{4} \times I_d \, , \quad (3.2)$$

Left eye boundary:

$$y_{\text{LB}} = y_l - \frac{1}{4} \times I_d \, , \quad (3.3)$$

Right eye boundary:

$$y_{\text{RB}} = y_r + \frac{1}{4} \times I_d \, , \quad (3.4)$$

The rest of the extraction process will be based on the defined eye region. Thus for the whole face region, $R_{\text{pf}}$, the width, $W_{\text{pf}}$, and height, $H_{\text{pf}}$, are determined as follows:
Table 3.1 Dimensions after preprocessing

<table>
<thead>
<tr>
<th>Subsection</th>
<th>Dimensions (pixels)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Whole face</td>
<td>30 × 30</td>
</tr>
<tr>
<td>Eyes</td>
<td>10 × 30</td>
</tr>
<tr>
<td>Eyes and Nose</td>
<td>20 × 30</td>
</tr>
<tr>
<td>Nose and Mouth</td>
<td>20 × 20</td>
</tr>
<tr>
<td>Mouth</td>
<td>10 × 30</td>
</tr>
</tbody>
</table>

\[ W_{pf} = H_{pf} = y_{RB} - y_{LB} \]  

The mouth region, \( R_{mo} \), is defined as the lower portion of the \( R_{pf} \) where the width, \( W_{mo} \), and height, \( H_{mo} \), can be approximated as follows:

\[ W_{mo} = W_{pf} \]  

\[ H_{mo} = \frac{1}{4} \times W_{pf} \]  

Following the notations defined earlier for each region, the upper face region (eyes and nose region), \( R_{up} \), and lower face region (nose and mouth region), \( R_{low} \), are approximated as follows:

\[ R_{up} = R_{pf} - R_{mo} \]  

\[ R_{low} = R_{pf} - R_{eye} \]  

Figure 3.8 shows the outcome of the face component extraction. Next, the image size of each subsection was reduced in order to decrease the processing time, as shown in Table 3.1. Lastly, a normalization process is applied to all the obtained subsections to reduce the effects of illumination. Figure 3.9 summarizes the preprocessing stage.

Training Stage

In the training stage, all the training images go through the preprocessing stage and each face image is divided into 5 subsections. Then, each subsection undergoes the FLD process and its respective space is created, i.e., Fisher-face, Fisher-eyes, Fisher-eyesnose, Fisher-nosemouth and Fisher-mouth space. The resulting weight distributions were obtained by mapping them onto their respective Fisher spaces for the recognition stage. After training, only weight distributions and the respective Fisher images from all five spaces are kept for recognition purpose. Figure 3.10 summarizes the training stage. Figure 3.12 shows the flow chart of the HFLD training stage.

Recognition Stage

In the recognition stage, first, the test image is divided into five subsections through the preprocessing stage. The dimensions of each subsection was reduced to the dimensions described in Table 3.1, followed by normalization. After mapping onto the respective Fisher spaces, the simple \( k \)-nearest neighbor (\( k \)-nnr) Euclidean distance classifier was utilized. In the face identification task, first, each hit scores a point. Then, the identity is matched with the users with the highest score or the users that has the shortest weight distribution distance when a majority decision could not be made (for example, a 2-2-1 vote). In the face authentication task, first, each hit scores a
point. Then, majority voting will decide whether the test image is from the training set or not, and if so, which class it belongs to. For example, if a test image has score higher than 3 for user A, then, the test image is said to be user A. Otherwise, the test image is from a unknown person (i.e., a intruder). **Figure 3.11** summarizes the testing stage. **3.13** shows the flow chart of the HFLD testing stage.
Figure 3.10 Overview of the HFLD training stage
Figure 3.11 Overview of the HFLD testing stage
Eye detection
Face component extraction
Image size reduction
Image normalization
Initiate $x = 1$

$x < 5$
Yes

Construct matrix $C_x$

Obtain overall mean, $\mu_{o,x}$ and class mean, $\mu_{i,x}$

Obtain within-class scatter, $S_{w,x}$, and between-class scatter matrices, $S_{b,x}$

Obtain $S_{ww,x} = V_{pca,x}^T S_{b,x} V_{pca,x}$ and $S_{bb,x} = V_{pca,x}^T S_{b,x} V_{pca,x}$

Obtain projection, $W_{opt,x}$

Obtain weight distribution, $d_x$

$x$ represents each face component as follows:
$x = 1$, face
$x = 2$, eyes
$x = 3$, eyesnose
$x = 4$, nosemouth
$x = 5$, mouth

Storing overall mean, $\mu_{o,x}$, fisher vectors, $W_{opt,x}$, and weight distribution, $d_x$

Remove all training images

Increment $x$

Figure 3.12 Algorithm flow of the HFLD training stage
Start

- Eye detection
- Face component extraction
- Image size reduction
- Image normalization
- Initiate $x = 1$

Decision: $x < 5$
- Yes: Obtain weight distribution, $d_x$
- No: Increment $x$

Comparison
- Obtain score ($x$)
- Increment $x$

Face Authentication
- Score $\geq 3$ for the same person?
- Yes: User
- No: Intruder

Face Identification
- Identity match with the person with highest score

Obtain weight distribution, $d_x$ from training stage

$x$ represents each face component as follows:
- $x = 1$, face
- $x = 2$, eyes
- $x = 3$, eyesnose
- $x = 4$, nosemouth
- $x = 5$, mouth

Figure 3.13 Algorithm flow of the HFLD testing stage
3.4 Experimental Results and Discussions

Experiments are conducted to evaluate the performance of the HFLD method in both face authentication and face identification task. False acceptance rate (FAR), false rejection rate (FRR), overall recognition rate (ORR) and identification rate (IR) are defined as follows:

\[
\text{FAR} = \frac{n_a}{N_u} \times 100\% ,
\]

\[
\text{FRR} = \frac{n_r}{N_k} \times 100\% ,
\]

\[
\text{ORR} = \left(1 - \left(\frac{n_a + n_r}{N_t}\right)\right) \times 100\% ,
\]

\[
\text{IR} = \frac{n_i}{N_t} \times 100\% ,
\]

where:

- \(n_a\): Number of falsely accepted intruders
- \(n_r\): Number of falsely rejected users
- \(n_i\): Number of correctly identified users
- \(N_u\): Number of intruder attempts
- \(N_k\): Number of user attempts
- \(N_t\): Total number of testing images

The performance of the proposed HFLD is evaluated in two face databases: (i) AT&T dataset and (ii) MZ dataset. The details regarding both these databases will be described in sections 3.4.1 and 3.4.2 respectively.

3.4.1 AT&T dataset

HFLD is tested using face images from the AT&T’s “dataset of Faces”† that was formerly known as “The Olivetti Research Laboratory (ORL) dataset of Faces”. The dataset consists of 400 face images (10 images for each individual for a total of 40 individuals) with variations in pose, illumination, facial expressions (open/closed eyes, smiling/non-smiling) and accessories (with glasses/without glasses). The frontal images have a tilt and rotation tolerance of up to 20° and a 10% scale tolerance. All the subjects were taken against a black background over a course of 2 years.

Figure 3.14 shows some face image examples from the AT&T dataset.

Experiment I: Optimal Configurations of the Proposed HFLD

The distance threshold, \(\theta\) defines the maximum allowable distance from a class, and is determined by the value of half the largest distance between two classes in the training set:

\[
\theta = \alpha \times \max\{\|\mathbf{d}_a - \mathbf{d}_b\|\}
\]

where \(a, b = 1, 2, ..., p\), for \(p\) number of classes.

†This dataset is developed by AT&T Laboratories Cambridge, (http://www.cl.cam.ac.uk).
Threshold values, $\theta$, in Eq. 3.14 for each subsection $x$ will affect the recognition rate (FRR, FAR, ORR). If we set the $\alpha$ value to 0, all testing images will assumed to be intruders, therefore $FRR = 100\%$ and $FAR = 0\%$. As the $\alpha$ value is gradually increased, FRR will be decreased because more testing images will be considered as users. At the same time, there will be an increasing number of intruders that are wrongly accepted as users from the training set. Figure 3.15 shows the effect of increasing the threshold value towards FAR and FRR.

Most of the face recognition researchers choose the threshold value when $FRR = FAR$. However, since the proposed HFLD is for security application, we need to choose threshold values that have low FAR to prevent intruders from penetrating through the security. Even though the threshold values will cause higher FRR, which consequently causes some users to be denied from accessing, the users can always re-attempt to access again.

The proposed HFLD requires five threshold values for the five subsections. The FAR versus FRR plot is obtained for each subsection and the $\alpha$ values for each subsections with the lowest possible FAR and resonable FRR are selected. To select good $\alpha$ values for the proposed HFLD, different combinations of the threshold values are experimented with, to obtain the highest ORR with lowest possible FAR. The optimized $\alpha$ values combination selected for HFLD as shown in Table 3.2.

Several minimum distance classifiers such as Euclidean, Mahalanobis and Soft-Decision clas-
Fisher-Eyes Detection

Adjust and obtain the eyes region

Obtain the face region

Obtain the mouth region

Obtain the eyes and nose region

Obtain the nose and mouth region

Figure 3.15 FAR and FRR versus different threshold values

Table 3.2 Optimum value of $\alpha$

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>$\alpha$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fisher-face</td>
<td>0.8</td>
</tr>
<tr>
<td>Fisher-eyes</td>
<td>0.8</td>
</tr>
<tr>
<td>Fisher-eyesnose</td>
<td>0.65</td>
</tr>
<tr>
<td>Fisher-nosemouth</td>
<td>0.7</td>
</tr>
<tr>
<td>Fisher-mouth</td>
<td>0.7</td>
</tr>
</tbody>
</table>

Fisher [64] were implemented for the proposed HFLD. Table 3.3 shows the results from the three different classifiers. It is clear that the simple Euclidean distance is the most suitable classifier. Even though the Mahalanobis classifier is reported to have excellent performance for PCA algorithms [65], it is not suitable for LDA-based algorithms such as HFLD. The LDA-based algorithm only keeps $p - 1$ eigenvectors and $p - 1$ Fishervectors for their respective Fishervectors, so if there are 25 classes in the training set, there are only 24 eigenvectors with 24 eigenvalues. This is unlike the PCA-based algorithm, where 25 classes with 8 training images each, will result in 200 eigenvectors and eigenvalues. Therefore, the eigenvalue strengths for PCA-based algorithms are closer to each other, hence, the impact of eigenvalues in Mahalanobis is not as extreme as that of the LDA-based algorithms. The Soft-Decision classifier uses eigenvalues as well, but the effect is reduced by the weights, therefore, it achieves better performance compared to the Mahalanobis classifier.

The majority voting score indicates the minimum number of votes required for the test image to be authenticated. For example, when the majority voting score is set to 3, and the score for a particular test image is $\geq 3$, then the test image is identified. If it is less than that, it is classified as an intruder. Table 3.4 shows the results of using several majority voting scores on the AT&T
Table 3.3 Comparison result of different classifier

<table>
<thead>
<tr>
<th>Minimum Distance Classifier</th>
<th>ORR</th>
<th>IR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Euclidean</td>
<td>99.2%</td>
<td>97.5%</td>
</tr>
<tr>
<td>Mahalanobis</td>
<td>60.0%</td>
<td>46.3%</td>
</tr>
<tr>
<td>Soft-Decision</td>
<td>85.6%</td>
<td>55.0%</td>
</tr>
</tbody>
</table>

Table 3.4 Comparison of results for different voting scores

<table>
<thead>
<tr>
<th>Voting Score</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Face Authentication</td>
<td>99.2%</td>
<td>86.4%</td>
<td>69.6%</td>
</tr>
</tbody>
</table>

dataset. The results show that the majority voting score of 3 yielded the best recognition rates.

Table 3.5 shows the effect of different values of $k$ for the $k$-nnr classifier when measurements are made between the weight distribution of the training set images and the test image. Three different values of $k$ were tested, i.e., 1, 3, and 5. When $k=1$, the test image belongs to the class with the shortest distance. When $k=3$, the 3 shortest distances will be considered when classifying the test image. In this case, if 2 of the shortest distances were from the same class, e.g., class A, the test image belongs to class A. If all of the 3 shortest distances are from different classes, the test image belongs to the class that has the minimum distance. Similarly, for $k=5$, if majority of the shortest distances were from a same class e.g., class B, the test image belongs to class B. Otherwise, if there is a tie for the number of shortest distances or if all of the shortest distances come from different classes, then the test image belongs to the class with the minimum distance. The results show that $k=1$ gives the best recognition rates.

**Experiment II: Face image authentication tests against PCA and LDA-based global approach algorithm**

The objective of Experiment II is to determine whether the test face image belongs to an user from the training set or an intruder using the proposed HFLD. 40-cross validation experiments are performed to verify the output result. 40 different seeds of random number generator for 40 different combinations of train and test images are used in each round of the experiment. In each experiment, we randomly select 25 individuals from the AT&T dataset, with 8 images for the training set (training) and the remaining 2 images as testing images. The face images of the 15 remaining individuals are used as face images of intruders, with 5 images per person.

The same experiment setup was repeated using other popular PCA and LDA-based global approach algorithm, i.e., direct linear discriminant analysis (DLDA) [43], Chen’s linear discriminant analysis algorithm (CLDA) [66], PCA [32], and FLD [37] to compare the performance measured in terms of FAR, FRR, ORR and computation time. All the approaches compared underwent the same optimization process as the proposed HFLD to obtain good $\alpha$ value for the threshold value during classification.

Table 3.6 shows the comparison of the averaged results for Experiment II of the proposed HFLD with DLDA, CLDA, PCA, and FLD respectively. From the table, it was observed that the proposed HFLD outperformed all the other algorithms in distinguishing users from intruders. With 0% FAR, the proposed HFLD successfully prevented all 75 intruders’ attempts from entering the home. On the contrary, DLDA, CLDA, PCA, and FLD have higher FARs. These global approach face recognition algorithms do not handle unknown testing images as well. Also, it can be observed that the threshold difference between users and intruders are not sufficient if
Table 3.5 Comparison results based on \textit{k-nnr} Classifier

<table>
<thead>
<tr>
<th>\textit{k-nnr}</th>
<th>\textit{k} = 1</th>
<th>\textit{k} = 3</th>
<th>\textit{k} = 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Face Authentication</td>
<td>99.2%</td>
<td>91.2%</td>
<td>96.8%</td>
</tr>
<tr>
<td>Face Identification</td>
<td>97.5%</td>
<td>80.0%</td>
<td>45.0%</td>
</tr>
</tbody>
</table>

Table 3.6 Comparisons of recognition rates, error rates and computation time for Experiment II

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>FAR</th>
<th>FRR</th>
<th>ORR</th>
<th>Training Time (s)</th>
<th>Recognition Time (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proposed HFLD</td>
<td>0.0%</td>
<td>2.0%</td>
<td>99.2%</td>
<td>14.2268</td>
<td>0.3606</td>
</tr>
<tr>
<td>DLDA [43]</td>
<td>5.3%</td>
<td>12.0%</td>
<td>92.0%</td>
<td>11.7018</td>
<td>0.0805</td>
</tr>
<tr>
<td>CLDA [66]</td>
<td>8.0%</td>
<td>14.0%</td>
<td>89.6%</td>
<td>21.5889</td>
<td>0.0814</td>
</tr>
<tr>
<td>PCA [32]</td>
<td>4.0%</td>
<td>24.0%</td>
<td>88.0%</td>
<td>1.5264</td>
<td>0.0840</td>
</tr>
<tr>
<td>FLD [37]</td>
<td>5.3%</td>
<td>36.0%</td>
<td>82.4%</td>
<td>4.6346</td>
<td>0.0778</td>
</tr>
</tbody>
</table>

only the whole face image is used as a test image.

The unique hybrid structure of the HFLD is able to verify that the inconsistency in the classification of testing images for the Fisher subsections i.e., Fisher-face, Fisher-eyes, Fisher-eyesnose, Fisher-nosemouth and Fisher-mouth is the indication of an intruder’s presence. HFLD yielded a significant 16.8\% improvement of ORR compared to conventional FLD. The FRR showed that 2\% of the users are denied access to the premises.

The last two columns of Table 3.6 shows the computation times for training and testing stage between the proposed HFLD with DLDA, CLDA, PCA, and FLD respectively. Even though the proposed method has an additional step of dividing the face into five components, the computation time is still relatively short, i.e., only 3 ms longer than the other PCA and LDA-based global algorithms.

Figure 3.16 shows the comparison results for proposed HFLD with other algorithms when \( y \) number of images per user are presented during the training stage. As expected, the accuracy in ORR is increasing with the increasing number of training images. Except during \( y = 2 \) when the proposed HFLD has the same ORR as some of the compared algorithms, it shows that the proposed HFLD is able to achieve significantly higher ORR compared to the other PCA and LDA-based global algorithms.

**Experiment III: Face image authentication tests against other hybrid approach algorithm**

Experiment III has the same experiment setup as Experiment II. However in this experiment, the proposed algorithm is compared with other hybrid approach algorithms, i.e., multimodel PCA (MMPCA) [55], sub-holistic PCA (SHPCA) [60], and modular subspace for LDA (MSLDA) [67]. These three algorithms are chosen to compare due to their unique way of dividing the face into subsections. All the hybrid approaches underwent the same optimization process as the proposed HFLD to obtain good \( \alpha \) value for the threshold value during classification.

Table 3.7 shows the comparison of the results for Experiment III of the proposed HFLD with MMPCA, SHPCA, and MSLDA respectively. It was clear that the proposed HFLD outperformed all the other hybrid approach algorithms in distinguishing users from intruders.

Unlike the other hybrid-based algorithm that assumed equal importance for all the subsections, MMPCA calculates the subsections through a weighting scheme of 6:5:4:3:2 for the pure face, brows and eyes, eyes, nose and mouth respectively. This method performed the worst overall compared to SHPCA and MSLDA. The high FAR rate shows that some face features such as eye
(without brows) and nose are not sufficient enough to create the boundary to distinguish users from intruders. A separate experiment was attempted with a weightless MMPCA by excluding the weighting scheme suggested in the work. The result of this weightless MMPCA is better compared to MMPCA with weighting scheme.

Even though SHPCA has some similarity with the proposed HFLD, the result shows that by dividing the face directly into four sections without considering the features is not able to reduce FAR percentage as well as the proposed HFLD. Depending on the head orientation, one forth of the face is sometimes unrecognizable.

Unlike the other hybrid approaches, MSLDA divides face into 3 subsections only, i.e., whole face, eyes and nose, and eyes. In the majority voting scores, the scores for a test image is less than 2 then it is classified as unknown person or intruder. The probability of errors occurred in two out of three is higher than three out of five. Therefore this method suffers the lowest ORR among the rest of the approaches that have 5 subsections.

**Experiment IV: Face image identification tests against PCA and LDA-based global approach algorithm**

The objective of Experiment IV is to identify each user within the training set using the proposed HFLD. Similar to the previous experiments, 40-cross validation experiments are performed to verify the output result. 40 different seeds of random number generator for 40 different combinations of train and test images are used in each round of the experiment. In each experiment, all 40 individuals were used, with 8 face images randomly chosen for each person for the training set.
Table 3.7 Comparisons of recognition rates, error rates and computation time for Experiment III

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>FAR</th>
<th>FRR</th>
<th>ORR</th>
<th>Training Time (s)</th>
<th>Recognition Time (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proposed HFLD</td>
<td>0.0%</td>
<td>2.0%</td>
<td>99.2%</td>
<td>14.2268</td>
<td>0.3606</td>
</tr>
<tr>
<td>MMPCA [55]</td>
<td>17.33%</td>
<td>2.0%</td>
<td>88.8%</td>
<td>8.9571</td>
<td>0.3929</td>
</tr>
<tr>
<td>MMPCA (weightless)</td>
<td>10.67%</td>
<td>2.0%</td>
<td>92.8%</td>
<td>8.9571</td>
<td>0.3813</td>
</tr>
<tr>
<td>SHPCA [60]</td>
<td>4.0%</td>
<td>6.0%</td>
<td>95.2%</td>
<td>6.7625</td>
<td>0.3401</td>
</tr>
<tr>
<td>MSLDA [67]</td>
<td>6.67%</td>
<td>16.0%</td>
<td>89.6%</td>
<td>5.5028</td>
<td>0.2708</td>
</tr>
</tbody>
</table>

Table 3.8 Comparison results for Experiment IV

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>IR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proposed HFLD</td>
<td>97.5%</td>
</tr>
<tr>
<td>DLDA</td>
<td>96.25%</td>
</tr>
<tr>
<td>PCA</td>
<td>95.0%</td>
</tr>
<tr>
<td>FLD</td>
<td>95.0%</td>
</tr>
<tr>
<td>CLDA</td>
<td>93.75%</td>
</tr>
</tbody>
</table>

Table 3.8 shows the comparison results for Experiment IV of the proposed HFLD with DLDA, CLDA, PCA, and FLD respectively. The table shows that the proposed HFLD was able to achieve the highest identification rate when compared with the other algorithms.

Experiment V: Face image identification tests against other hybrid approach algorithm

Experiment V has the same experiment setup as Experiment IV. However, in this experiment, the proposed algorithm is compared with other hybrid approach algorithms, i.e., SHPCA, MMPCA and MSLDA.

Table 3.9 shows the comparison of the results for Experiment V of the proposed HFLD with SHPCA, MMPCA, and MSLDA, respectively. The proposed HFLD is able to achieve the highest identification rate among the other hybrid approaches. The proposed hybrid structure is able to increase the confidence level of the user identification task with the better face features selected in each subsection.
Table 3.9 Comparison results for Experiment V

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>IR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proposed HFLD</td>
<td>97.5%</td>
</tr>
<tr>
<td>SHPCA</td>
<td>96.3%</td>
</tr>
<tr>
<td>MMPCA</td>
<td>95.0%</td>
</tr>
<tr>
<td>MSLDA</td>
<td>92.5%</td>
</tr>
</tbody>
</table>

3.4.2 MZ dataset

We created another dataset henceforth referred to as the MZ dataset to simulate the scenario of a typical Malaysian household. The dataset consists of 8 people from one family and 13 other individuals to act as intruders, overall consisting of males and females of Chinese, Indian and Malay origins, their ages ranging from 16 to 76 years old. The dataset consists of a total of 420 frontal face images (20 images for each individual for a total of 21 individuals). The individual face images were taken with variations in pose, illumination, facial expression (open/closed eyes, smiling/non-smiling) and accessories (with glasses/without glasses). The images were taken against different backgrounds (indoor/outdoor) and lightings (natural light/fluorescent light). All the individual were asked to be in an upright, frontal position. The images were cropped to 92 × 112 pixels, with 8-bits 256 grey levels per pixel. Figure 3.18 shows some face image examples from the MZ dataset.
Experiment VI: Face image authentication and identification test with MZ dataset

Experiment VI has the same experiment setup with Experiment II and IV but with the MZ dataset. For the face authentication test, in each experiment, 8 family members from the dataset are selected for the training set with 10 images for training and the remaining 10 images as testing images. The face images of the 13 remaining individuals are used as face images of intruders, with 10 images per person. For the face identification test, in each experiment, 8 family members from the dataset are selected to be in the training set with 10 images for training and the remaining 10 images as testing images.

Table 3.10 shows the results for the proposed HFLD in Experiment VI. The results showed that the proposed HFLD algorithm is able to achieve an ORR of 98.6% with an FAR of 0% and an FRR of 3.8%. This is implying that none of the 130 face images of intruders was falsely accepted even though some household members were denied access. Hence with this algorithm, one can be sure that the home security system is able to prevent intruders from entering the secured premises. It showed even though the significantly increased number of intruders attempts, the proposed HFLD is still able to achieve good ORR.

The result also showed that the proposed HFLD achieved an identification rate of 96.3%. Even though it is slightly lower than the recognition rates achieved with the AT&T dataset, it is still considered quite reliable as the MZ dataset contains four siblings who look relatively alike.
Table 3.10 Comparison of results for Experiment VI and Experiment VII

<table>
<thead>
<tr>
<th>Experiment</th>
<th>FAR</th>
<th>FRR</th>
<th>ORR</th>
<th>IR</th>
</tr>
</thead>
<tbody>
<tr>
<td>IV</td>
<td>0.0%</td>
<td>3.8%</td>
<td>98.6%</td>
<td>-</td>
</tr>
<tr>
<td>V</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>96.3%</td>
</tr>
</tbody>
</table>

Again, the HFLD managed to achieve an excellent identification rate that is higher than 95%.

3.5 Conclusions

A face recognition system using a HFLD analysis method that is able to identify each user as well as authenticate between users and intruders was proposed to examine the performance of the basic linear-based approaches in the exact measurement face classification. The proposed HFLD methodology compensates for the low performance of the conventional FLD when tested with face images from intruders by combining the benefits of both global and local approaches. The proposed method divides the input face image into five subsections: whole face, eyes, eyes and nose, nose and mouth, and mouth. Each subsection will be projected onto its own FLD space. By checking the inconsistency of the decision from each subsection’s FLD space, the final decision will be obtained through a voting system. The proposed methodology is able to utilize the efficiency of FLD algorithm in face identification application, yet improving its recognition rate for face authentication application when handling face images from intruders.

Through the experiments, the proposed hybrid linear-based method outperformed all the other global linear-based approach conventional methods and all other hybrid linear-based methods that was compared. The proposed method achieved 99.2% and 98.6% in overall recognition rate (ORR) in face authentication for the AT&T and MZ databases, respectively. Also, this method yielded 97.5% and 96.3% in identification rate (IR) for face identification for the AT&T and MZ databases, respectively. It was also clear that the majority voting score of 3 and the 1-nnr Euclidean distance classifier yielded the optimum configuration for the proposed methodology.
Chapter 4

Abstract Measurement
Face Classification

This chapter presents an overview for one of the abstract measurement face classification, i.e., face classification based on personal attractive preference to challenge the performance of the linear-based algorithm that we propose. This chapter begins with an introduction. Then, it describes the issues with the conventional approaches that motivates us to propose the class-specific eigenfaces reconstruction methodology. Experiments are carried out to verify the relative performance of the proposed method with respect to the conventional PCA-based methods.
4.1 Introduction

The abstract measurement face classification is defined as an approach that classifies faces into predefined classes with vague and ambiguous definition. This approach is relatively new and therefore have a lot of room for improvement unlike the exact measurement face classification. In this thesis, we will study the abstract measurement face classification based on personal attractive preference.

Face classification based on attractiveness can be view as a process of categorizing faces into several level of attractiveness based on the subjective view of how appealing that face is towards an individual or based on the general opinions of a group of individuals. This process has potentially profit generating applications. For example, online match making systems, and automatic face image filtering systems for beauty contests. Among the examples of the applications mentioned, the application systems are required to able to classify the faces based on several customized level of attractiveness depending on the user. Furthermore, the attractive level may also be defined based on some facial characteristic. Hence, the application can also be utilized for the talent castings for movies, drama, commercials and fashion show. With this, there should not be any gender and ethnicity restrictions. It should be able to handles faces from both genders and different ethnicity backgrounds, i.e., *non gender specific* and *non ethnicity specific*.

Even though it is shown that the FLD performed better than PCA in Chapter 3, this approach is not suitable for application with only very few classes because the between class-variation is too low [37]. For instance, for an application with only 3 classes, there are only 2 meaningful Fisherfaces and they will not able to describe all the face images within the dataset well. Therefore, the PCA method [31] appears to be an attractive approach because it is simple and easy to implement, does not require manual labeling on face features and prior knowledge (i.e., gender and ethnicity) of the faces and suitable for applications that are classifying faces into only a few classes. The conventional steps of using PCA can be summarized as follows: i) all the faces from different classes are grouped together as one huge training set, ii) PCA is applied on the training set for dimensionality reduction, iii) weights to represent each training image are obtained, and iv) comparison is made between the weights of the test image and training images using some classification methods such as KNN and SVM.

Although the conventional steps for PCA showed tremendous success with those exact measurement face classification applications, this conventional practice is not a good way to utilize PCA for in this application. The practice of comparing the weights between the test image and training images may be interpreted as a process to locate a single face among the training images that is visually the closest match to the test image. However, the objective is to classify faces between the classes, i.e., *attractive* (A), *common* (C), *non attractive* (N) and not between individuals, therefore this motivates us to propose another methodology to utilize PCA.

Conventional PCA-based methods produced eigenfaces that contain overall face information from all the classes. The general eigenfaces produced by training face images from all the classes together are not distinct enough to represent all the specific face characteristic information from each class. To solve this issue, in this work, we propose a new methodology that trains each class separately to produce class-specific eigenfaces that is unique to each class. These eigenfaces are able to represent each class distinctly compared to those produced through the conventional methods. Because eigenfaces from each class span different spaces, we cannot compare the weights between the test images and training images directly. Therefore, we propose to classify test images in the spatial space (pixel space) by comparing the original test image with reconstructed images using the class-specific eigenfaces produced by the proposed method.

Also, conventional approaches are evaluated based on the general perspective of face attrac-
tiveness. They collected opinions from a pool of participants and pre-classified faces into different classes based on those overall opinions. Thus, this does not represent the opinion of each participant and therefore, unable to verify the performance for applications such as online match making system. Therefore, we evaluate the proposed method and other conventional PCA-based approaches by their performance to classify faces that have been pre-classified into 3 classes, i.e., attractive (A), common (C), non attractive by each participants.

The proposed method is non gender specific and non ethnicity specific and is able to handle faces from different gender and ethnicity by changing the training and test images, but since all the participants available for this research are of Asian ethnicity, also, because cross race faces are generally more difficult to recognize and judge when compared to faces from own race [68], we focus only on the Asian faces. Furthermore, to the best of our knowledge, only the attractiveness of female faces has been evaluated so far [19, 20, 21, 22, 23] but we evaluate the proposed face classification based on personal attractive preference using class-specific eigenfaces reconstruction method for both female and male faces.

4.2 Background

Earlier work in face classification based on attractiveness was first introduced by Eisenthal et al. [19] using Principle Component Analysis (PCA) [31]. In their approach, face images are represented in 2 ways: (i) configuration representation and (ii) featural representation. Configuration representation is raw grayscale pixel values of the face images, denoted as pixel images, whereas featural representation is the manual measurements of 37 facial feature distances and ratios that reflect the geometry of the face, denoted as feature vectors. Point-biserial correlation technique is used to select the most relevant eigenfaces produced from PCA while K-Nearest Neighbors (KNN) [69] and Support Vector Machines (SVM) [70] are used as classifier. Although the pixel images are non gender and non ethnicity specific, the feature vectors are designed specifically for female faces only. For example, the thickness of the eyebrow might be important to determine the attractiveness of female faces but it is not as important for male.

Turkmen et al. [20] extended the work by introducing an ellipse face mask to mask off and remove distractions such as hair and other accessories for improvement in accuracy. The top 60 % eigenfaces with the largest eigenvalues are selected and SVM is used as classifier. Although the pixel images are non gender and non ethnicity specific, the feature vectors are designed specifically for female faces only. For example, the thickness of the eyebrow might be important to determine the attractiveness of female faces but it is not as important for male.

Mao et al. [21] proposed a four categories (beautiful, good, common and poor) classifier of Chinese female facial beauty using SVM. In this approach, 17 geometric measurements of the face that are manually obtained by using specific vertical and horizontal lines are used for face representation. The specific vertical and horizontal lines need to be modified for male and other ethnicity.

While the works of Eisenthal et al. and Mao et al. involved landmarks, i.e., manual annotation feature points on face images, Gray et al. [22] proposed a method to predict facial beauty without manual labeling. In the research, a complex, biologically-inspired model with multiple layers of trainable feature extraction is proposed. The features are then classified using Hubel-Wiesel Neural Network.

Kanghae et al.[23] used face attractiveness as one of the features in face recognition application. The face attractiveness is measured by the distances between the face feature and the golden mask. Different golden masks are required for different gender and different ethnicity group. The system is trained and classified using a neural network model.
4.3 Methodology

This section discusses the concept of the proposed method. Then, it presents the overview of the proposed method.

4.3.1 Conceptual differences from the conventional PCA-based method

The main difference between the PCA-based conventional method in face attractiveness classification\cite{20, 19} and the proposed method during training stage is highlighted in Figure 4.1. During the training stage, the conventional method combine face images from all the 3 classes, i.e., attractive (A), common (C), non attractive (N), into a huge training set. Then, PCA is used to produce eigenfaces, $U$, that contain general face information from all the 3 classes. On the contrary, we propose the usage of PCA to produce eigenfaces for each class separately. The class-specific eigenfaces produced, $U_A$, $U_C$, and $U_N$, each contains unique face information that is specific to its respective class.

Also, the main difference between the PCA-based conventional method in face attractiveness classification\cite{20, 19} and the proposed method during the testing stage is highlighted in Figure 4.2. The conventional methods perform the classification in the weights space, $\Omega$, i.e., in the projected subspace of original images spanned by each columns of $U$. On the other hand, the new methodology propose to perform classification in the spatial space (pixel space).

Figure 4.3 and Figure 4.4 illustrate the eigenfaces obtained from (a) the conventional PCA-based method, and (b-d) the proposed method for (b) attractive, (c) common and (d) non attractive class, respectively. These eigenfaces are more specific to each respective class compared to those produced by the conventional method.

In the proposed method, to obtain the class that contains the most similar face features of the test image, we first reconstruct test image for each class using Eq. 2.4 from class-specific eigenfaces $U_A$, $U_C$ and $U_N$. The reconstructed images will contain the face information from each class that best represent the test image. Then, the test image and the reconstructed images are compared in the spatial space (pixel space) to produced similarity values $\delta_A$, $\delta_C$ and $\delta_N$. As an example, assuming a test image is from class A, since the class-specific eigenfaces $U_A$ contain specific face information of that class, the class A reconstructed test image will look more similar to the test image compared to the reconstructed image using other eigenfaces. The test image be classify as class A after measuring and comparing the similarity value between the reconstructed images from all the classes.

The flow of the conventional PCA-based face attractiveness classification method is summarized in Figure 4.5. The conventional PCA-based method combines the sets of the attractive faces, $X_A$, common faces, $X_C$ and non attractive faces, $X_N$ into a single training set, $X_{tr}$. The PCA is used to produce eigenfaces, $U_{tr}$, from the training set, $X_{tr}$. The weights of the training images, $\Omega_{tr}$, that consist of face features of face images from all the classes are obtained using Eq. 2.5. In the testing stage, test image, $x$ is also mapped to the same space to obtain weights, $\Omega$. Then, a classifier such as SVM and KNN is performed in the weight space.
\( \Omega_a = a_1U_1 + a_2U_2 + \cdots + a_MU_M \)
\( \Omega_b = b_1U_1 + b_2U_2 + \cdots + b_MU_M \)
\( \Omega_c = c_1U_1 + c_2U_2 + \cdots + c_MU_M \)
\( \Omega_m = m_1U_1 + m_2U_2 + \cdots + m_MU_M \)

(a) Conventional method  
(b) Proposed method

**Figure 4.1** Conventional method vs. proposed method during training stage
(a) Conventional method

Test Image

Eigenfaces, $U_A$

Reconstruct

Compare

$\Omega_a = a_1U_1 + a_2U_2 + \ldots + a_MU_M$

(b) Proposed method

Test Image

Eigenfaces, $U_C$

Reconstruct

Compare

$\Omega_c = c_1U_1 + c_2U_2 + \ldots + c_MU_M$

Eigenfaces, $U_N$

Reconstruct

Compare

$\Omega_m = m_1U_1 + m_2U_2 + \ldots + m_MU_M$

$\Omega_t = t_1U_1 + t_2U_2 + \ldots + t_MU_M$

Figure 4.2. Conventional method vs. proposed method during testing stage

$\delta_A$

$\delta_C$

$\delta_M$
Figure 4.3 Examples of eigenfaces for female dataset (a) Eigenfaces produced from conventional method (b) Attractive class eigenfaces produced from proposed method (c) Common class eigenfaces produced from proposed method (d) Non-attractive class eigenfaces produced from proposed method.
Figure 4.4 Examples of eigenfaces for male dataset (a) Eigenfaces produced from conventional method (b) Attractive class eigenfaces produced from proposed method (c) Common class eigenfaces produced from proposed method (d) Non-attractive class eigenfaces produced from proposed method
Figure 4.5 Block diagram of the conventional PCA-based face attractiveness classifier
4.3.2 Proposed Methodology

The flow of the propose methodology is summarized in Figure 4.6. The proposed method produces class-specific eigenfaces for three distinct classes, i.e., attractive (A), common (C), non attractive (N) that had been pre-classified by a participant. Face images from each class undergoes PCA separately using training set $X_A$, $X_C$ and $X_N$, respectively. Eigenfaces $U_A$, $U_C$, $U_N$, have the characteristic of the face images that had been classified by a participant as attractive, common and non attractive, respectively. Then, we reconstruct test image $x$ for each class, $x_A$, $x_C$ and $x_N$ through Eq. 2.4 using eigenfaces from each class $U_A$, $U_C$ and $U_N$, respectively.

Some examples of reconstructed images using eigenfaces from each class in the female and male dataset are shown in Figure 4.7 and Figure 4.8, respectively. Consider the top left set of images in Figure 4.7, it shows a test image, image (a), from female dataset that is pre-classified as attractive by Participant 1 and reconstructed images, image (b-d), using eigenfaces $U_A$, $U_C$ and $U_N$, respectively. Visually, the reconstructed image from attractive class, image (b), looks the most similar to original image (a) when compares to image (c) and image (d). Therefore, the test image is classify as attractive class. Similar arguments can be apply in the other examples in Figure 4.7 and Figure 4.8.

To determine the similarity values for each class, $\delta_A$, $\delta_C$, and $\delta_N$, between $x$ and $x_A$, $x_C$, and $x_N$, respectively, we consider the Euclidean distance to dictate which reconstructed image, i.e., $x_A$, $x_C$, and $x_N$ is closer to the original image, $x$. Let us denote the source and target images as $S$ and $T$, where $i \in [1, n]$ and its similarity value $\delta$ is computed as follows:

$$\delta = ||S - T||_2 = \sqrt{\sum_{i=1}^{n} [S_i - T_i]^2},$$  \hspace{1cm} (4.1)

where $S(i)$ and $T(i)$ denote the $i$-th components of $S$ and $T$, respectively. The following equation determines the membership of a given test image $x$:

$$V = \arg\min_{j \in A,C,N} \delta_j,$$  \hspace{1cm} (4.2)

where $j$ denotes the attractive (A), common (C), and non attractive (N) class.
Notation:

$x$ : Test image
$X_{tr}$ : Training set
$X_A$ : Reconstructed image from class A
$X_C$ : Reconstructed image from class C
$X_N$ : Reconstructed image from class N
$U$ : Eigenface
A : Attractive
C : Common
N : Non attractive

Class($x$) = \arg \min_{i \in \{A,C,N\}} \{\delta_i\}

Figure 4.6 Block diagram of the proposed method
**Figure 4.7** Examples of reconstructed images for female dataset. Original image classified as attractive (A), common (C) and non attractive (N) (from top to bottom in that order) by Participant 1 (a) Original Image (b) Reconstructed image using eigenfaces from attractive class (c) Reconstructed image using eigenfaces from common class (d) Reconstructed image using eigenfaces from non attractive class
Figure 4.8 Examples of reconstructed images for male dataset. Original image classified as attractive (A), common (C) and non attractive (N) (from top to bottom in that order) by Participant 1 (a) Original Image (b) Reconstructed image using eigenfaces from attractive class (c) Reconstructed image using eigenfaces from common class. (d) Reconstructed image using eigenfaces from non attractive class
4.4 Dataset and Experimental Setup

The dataset used in the experiment is a combination of 129 face images from the CAS-PEAL face dataset [71] and 351 face images collected from the internet. The dataset consists of 240 face images of Asian female and 240 face images of Asian male, a total of 480 face images. All the images are aligned and rotated according to the position of the iris. The face images are also cropped to remove unnecessary distraction such as hair and other accessories. Then, the images are converted into grayscale images and resized to $180 \times 240$ pixels. Some examples of face images in the female and male dataset * are illustrated in Figure 4.9 and Figure 4.10.

Face images in the dataset is pre-classified into 3 classes, i.e., attractive (A), common (C), and non attractive (N) by 7 different participants (4 female participants and 3 male participants), i.e., $P_1, P_2, ..., P_7$, based on their own personal attractive preference. The participants are instructed to classify faces that is attractive to them into the attractive (A) class, the faces that is non attractive to them into the non attractive (N) class, and those ambiguous faces into common (C) class. Here, the number of face images in each class is restricted to $n = \min(n_\alpha)$ for $\alpha \in \{A, C, N\}$ because each class consists of different numbers of faces. When $n_\alpha > n$, $n$ images are randomly selected from class $\alpha$.

40-cross validation experiments are performed to verify the output result. 40 different seeds of random number generator for 40 different combinations of train and test images are used in each round of the experiment. In each experiment, 50% of the available face images in each class are used as test images and the remaining as training images.

![Figure 4.9 Examples of face images in the female dataset](image)

![Figure 4.10 Examples of face images in the male dataset](image)

In this paper, PCA-based Eisenthal [19] and Turkmen [20] methods are used for comparison purposes to evaluate the performance of the proposed method. The training set used in the experiments for the proposed method and Eisenthal pixel image method are the same. However, ellipse face mask is applied on each of the face images in the training set for the Turkmen method.

*Face images in this dataset is used solely for research purpose only.

48
### Table 4.1 Algorithm List

<table>
<thead>
<tr>
<th>Color Code</th>
<th>Algorithm</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>Proposed Method</td>
</tr>
<tr>
<td>□</td>
<td>2</td>
<td>Eisenthal (Pixel) KNN [19]</td>
</tr>
<tr>
<td>□</td>
<td>3</td>
<td>Eisenthal (Pixel) SVM [19]</td>
</tr>
<tr>
<td>□</td>
<td>4</td>
<td>Eisenthal (Feature) KNN [19]</td>
</tr>
<tr>
<td>□</td>
<td>5</td>
<td>Eisenthal (Feature) SVM [19]</td>
</tr>
<tr>
<td>□</td>
<td>6</td>
<td>Turkmen KNN [20]</td>
</tr>
<tr>
<td>□</td>
<td>7</td>
<td>Turkmen SVM [20]</td>
</tr>
</tbody>
</table>

Also, the performance of the Eisenthal feature vector method is included for comparison purposes. Linear SVM function from LIBSVM [72] in Weka [73] is used as the SVM classifier in the experiments.

The computational order and computational time in the experiments are obtained through the female dataset experiment by Participant 1 with Euclidean distance similarity measurement, with 111 training images and 111 test images (37 images in each class for both training and testing set). The computational time is measured by using Dell XPS 8300, with Intel® Core™ i7-2600 CPU at 3.40GHz, 12.0 Gb RAM and 64-bit Windows 7 Ultimate operating system. The algorithms are coded in Matlab student version 7.12.0.635 (R2011a) 32-bit (win32). The time obtained are the average of the timing output repeated 40 times.

### 4.5 Experimental Results and Discussions

In this paper, we conduct 2 different sets of experiments: 3-class classification experiments to evaluate the performance of the proposed method to classify face images into attractive (A), common (C), and non attractive (N) class and 2-class classification experiments to evaluate the performance to classify only 2 classes that are less ambiguous, i.e., attractive (A) and non attractive (N) class.

#### 4.5.1 Three-class classification

The boxplots for the 40-cross validation experiments for the proposed method and other conventional methods for comparison purposes using female and male dataset are shown in Figure 4.11 and Figure 4.12, respectively. Y-axis shows the accuracy \( A_n \) of correctly classified faces for participant \( P_n \) as follows:

\[
A_n = \frac{\text{Sum of correctly classified images}}{\text{Sum of total of testing images}} \times 100\% \quad (4.3)
\]

X-axis shows the algorithms to be compared as listed in Table 4.1. The bottom and top of each box in the boxplot indicate the first and third quartiles, respectively. The line in the middle indicates the median of the accuracy from the experiments for each algorithm. The crosses are outliers from the experiments.

Due to the nature of the results obtained from the 40 experiments, i.e., 1) results obtained are not normally distributed, and 2) the results are related samples as the same train and test face
images are used for each experiment for all the algorithms, therefore, Wilcoxon signed-rank test [74] is selected as a statistical test to verify the obtained results as it is a non-parametric statistical hypothesis test used when comparing 2 related samples for non normalized distributed data.

While the results shown in Figure 4.11 and Figure 4.12 vary depending on the participants, the proposed method outperformed all the conventional methods for all the participants, with the confidence level of 95% according to the Wilcoxon signed-rank test. The proposed method also showed less variance in the output compared with other methods. This indicates a higher consistency in the results by the proposed method. For example, in Figure 4.11 for the female dataset, comparing the boxplots for participant 1 and participant 7, the proposed method showed relatively consistent in the size of the blue box indicating the results were more consistent. On the contrary, for the Eisenthal pixel image SVM method, i.e., algorithm 3, there are a very small variance in accuracy for participant 1 but large variance for participant 7. Comparing the boxplots for female and male dataset in Figure 4.11 and Figure 4.12, the results for female dataset are slightly better compared to the male dataset for this experiment setup. This might be because the female faces were easier for these participants to classify compared to male faces.

The averaged improvement in accuracy, \( \bar{A}_q \), of the proposed method when compared to conventional methods considered are shown in Table 4.2. The averaged improvement in accuracy between the proposed method \( p \) and other methods \( q \) for all 7 participants \( (P_n) \) among the 40 conducted experiments is obtained by the following equation:

\[
\bar{A}_q = \frac{1}{40 \times 7} \sum_{n=1}^{7} (A_{n,p} - A_{n,q}) .
\]  

(4.4)

In general, the proposed method showed noticable improvement compared with other conventional methods. In the female dataset, the Eisenthal pixel image based method with SVM performed the best yet the algorithm performed the worst with KNN classifier among the conventional methods. Turkmen method performed better when compared to Eisenthal method with KNN classifier, but Eisenthal method performed better when compared to Turkmen Method with SVM classifier. For the male dataset, as predicted, the Eisenthal feature based method performed the worst in both KNN and SVM classifier. This method is designed specially for female faces only. Some of features in the method such as the thickness of the eyebrow is only important to determine the attractiveness of female faces but it is not as important for male.

The confusion matrices for both dataset to indicate the performance of all the algorithms in each class for Participant 1 are shown in Figure 4.13 and Figure 4.14. The columns show predicted class and the rows indicate the known class. The values in the confusion matrices are the averaged percentage of each occurrence. In the confusion matrices, the proposed method showed a relatively high accuracy when compared to the conventional methods for all 3 classes. The proposed method also showed low false positive between the extreme class, i.e., attractive and non attractive class.

Overall, in the female dataset, the percentage of true positive in the non attractive class is higher than the other 2 classes for all the algorithms except for the Turkmen SVM method. The method had the most false positive and also true positive in the common class. This shows that the faces in the non attractive class for female are more distinct and therefore easier to classify compared to the other 2 classes. However, in the male dataset, the percentage of true positive in the attractive class is higher than the other 2 classes for most of the algorithms. This shows that as opposed to the female faces, the features for attractive male faces are more distinct and easier to classify compared to the other 2 classes. Many false positives occur in the common class. The faces in the common class are ambiguous faces that are considered as neither attractive nor non
attractive. These faces are likely to have less unique features as compared to the faces in attractive and non attractive class.
Figure 4.11 Results for 3-class female classification

Figure 4.12 Results for 3-class male classification
<table>
<thead>
<tr>
<th>Compared Algorithm (q)</th>
<th>Female</th>
<th>Male</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\bar{A}_q$ (%)</td>
<td>Standard Deviation</td>
</tr>
<tr>
<td>Eisenthal (Pixel) KNN</td>
<td>15.1</td>
<td>6.2</td>
</tr>
<tr>
<td>Eisenthal (Pixel) SVM</td>
<td>7.7</td>
<td>6.3</td>
</tr>
<tr>
<td>Eisenthal (Feature) KNN</td>
<td>11.7</td>
<td>4.0</td>
</tr>
<tr>
<td>Eisenthal (Feature) SVM</td>
<td>13.8</td>
<td>3.9</td>
</tr>
<tr>
<td>Turkmen KNN</td>
<td>10.8</td>
<td>2.7</td>
</tr>
<tr>
<td>Turkmen SVM</td>
<td>14.1</td>
<td>3.3</td>
</tr>
</tbody>
</table>
### Figure 4.13 Confusion matrix for female dataset

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>C</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>51.9</td>
<td>45.4</td>
<td>2.7</td>
</tr>
<tr>
<td>C</td>
<td>33.5</td>
<td>54.6</td>
<td>11.9</td>
</tr>
<tr>
<td>N</td>
<td>13.0</td>
<td>14.6</td>
<td>72.4</td>
</tr>
</tbody>
</table>

(a) Proposed Method

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>C</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>27.0</td>
<td>10.3</td>
<td>62.7</td>
</tr>
<tr>
<td>C</td>
<td>13.0</td>
<td>9.2</td>
<td>77.9</td>
</tr>
<tr>
<td>N</td>
<td>3.8</td>
<td>3.8</td>
<td>92.4</td>
</tr>
</tbody>
</table>

(b) Eisenthal (Pixel) KNN

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>C</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>45.4</td>
<td>15.7</td>
<td>38.9</td>
</tr>
<tr>
<td>C</td>
<td>24.8</td>
<td>14.8</td>
<td>47.6</td>
</tr>
<tr>
<td>N</td>
<td>12.4</td>
<td>6.5</td>
<td>86.5</td>
</tr>
</tbody>
</table>

(c) Eisenthal (Pixel) SVM

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>C</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>52.4</td>
<td>31.9</td>
<td>15.7</td>
</tr>
<tr>
<td>C</td>
<td>42.7</td>
<td>35.7</td>
<td>21.6</td>
</tr>
<tr>
<td>N</td>
<td>17.8</td>
<td>23.2</td>
<td>58.9</td>
</tr>
</tbody>
</table>

(d) Eisenthal (Feature) KNN

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>C</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>58.4</td>
<td>34.1</td>
<td>7.6</td>
</tr>
<tr>
<td>C</td>
<td>36.8</td>
<td>36.8</td>
<td>23.8</td>
</tr>
<tr>
<td>N</td>
<td>16.8</td>
<td>17.8</td>
<td>65.4</td>
</tr>
</tbody>
</table>

(e) Eisenthal (Feature) SVM

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>C</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>52.9</td>
<td>33.0</td>
<td>14.1</td>
</tr>
<tr>
<td>C</td>
<td>43.2</td>
<td>31.4</td>
<td>25.4</td>
</tr>
<tr>
<td>N</td>
<td>22.2</td>
<td>21.1</td>
<td>56.8</td>
</tr>
</tbody>
</table>

(f) Turkmen KNN

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>C</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>38.4</td>
<td>55.7</td>
<td>7.0</td>
</tr>
<tr>
<td>C</td>
<td>23.2</td>
<td>68.6</td>
<td>5.4</td>
</tr>
<tr>
<td>N</td>
<td>14.6</td>
<td>51.9</td>
<td>33.5</td>
</tr>
</tbody>
</table>

(g) Turkmen SVM
<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>C</th>
<th>N</th>
<th></th>
<th>A</th>
<th>C</th>
<th>N</th>
<th></th>
<th>A</th>
<th>C</th>
<th>N</th>
<th></th>
<th>A</th>
<th>C</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>72.3</td>
<td>16.1</td>
<td>11.6</td>
<td></td>
<td>50.3</td>
<td>32.9</td>
<td>16.8</td>
<td></td>
<td>52.9</td>
<td>29.7</td>
<td>17.4</td>
<td></td>
<td>53.5</td>
<td>29.7</td>
<td>16.8</td>
</tr>
<tr>
<td></td>
<td>32.9</td>
<td>46.5</td>
<td>20.6</td>
<td></td>
<td>31.0</td>
<td>32.9</td>
<td>36.1</td>
<td></td>
<td>26.5</td>
<td>40</td>
<td>30.3</td>
<td></td>
<td>33.5</td>
<td>32.9</td>
<td>33.5</td>
</tr>
<tr>
<td></td>
<td>17.4</td>
<td>35.5</td>
<td>47.1</td>
<td></td>
<td>9.7</td>
<td>46.5</td>
<td>43.9</td>
<td></td>
<td>9.0</td>
<td>27.1</td>
<td>63.9</td>
<td></td>
<td>20.7</td>
<td>46.2</td>
<td>34.2</td>
</tr>
</tbody>
</table>

(a) Proposed Method  
(b) Eisenthal (Pixel) KNN  
(c) Eisenthal (Pixel) SVM  
(d) Eisenthal (Feature) KNN  
(e) Eisenthal (Feature) SVM  
(f) Turkmen KNN  
(g) Turkmen SVM

Figure 4.14 Confusion matrix for male dataset
4.5.2 Extreme case: Two class classification

In this experiment, we only consider the 2 extreme classes, i.e., attractive (A) and non-attractive (N) by excluding the common (C) class. Faces in these classes are less ambiguous as compared to the faces in the common class. The boxplots for the 40-cross validation experiments for the proposed method and other conventional methods for comparison in 2-class classification for female and male dataset are shown in Figure 4.15 and Figure 4.16, respectively. Similar to 3-class classification experiments, the results in 2-class classification experiments also vary depending on the participants, the proposed method outperformed all the conventional methods for all the participants, with the confidence level of 95% according to the Wilcoxon signed-rank test.

As predicted, the accuracy for all the algorithms increased significantly when compared to the 3-class classification. Table 4.3 shows the averaged improvement in accuracy of the proposed method when compared to other conventional methods in the 2-class classification experiment. As opposed to the results for the 3-class classification experiments, Turkmen method showed better performance when compared to Eisenthal pixel image based method in 2-class classification. The inconsistency among the performance of the conventional methods in the 2-class and 3-class classification results showed that the methods are less robust as the results vary when the train and test images are different. Eisenthal feature vector method remains to be the lowest accuracy method among all in the male dataset.

<table>
<thead>
<tr>
<th>Compared Algorithm (q)</th>
<th>Female</th>
<th></th>
<th>Male</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\bar{A}_q$ (%)</td>
<td>Standard Deviation</td>
<td>$\bar{A}_q$ (%)</td>
<td>Standard Deviation</td>
</tr>
<tr>
<td>Eisenthal (Pixel) KNN</td>
<td>17.7</td>
<td>8.2</td>
<td>16.6</td>
<td>5.4</td>
</tr>
<tr>
<td>Eisenthal (Pixel) SVM</td>
<td>10.3</td>
<td>5.0</td>
<td>7.4</td>
<td>5.8</td>
</tr>
<tr>
<td>Eisenthal (Feature) KNN</td>
<td>9.1</td>
<td>3.2</td>
<td>26.5</td>
<td>6.1</td>
</tr>
<tr>
<td>Eisenthal (Feature) KNN</td>
<td>14.3</td>
<td>5.2</td>
<td>12.3</td>
<td>4.5</td>
</tr>
<tr>
<td>Turkmen KNN</td>
<td>6.3</td>
<td>3.4</td>
<td>11.0</td>
<td>3.2</td>
</tr>
<tr>
<td>Turkmen SVM</td>
<td>7.1</td>
<td>4.3</td>
<td>5.2</td>
<td>2.6</td>
</tr>
</tbody>
</table>
Figure 4.15 Results for 2-class female classification

Figure 4.16 Results for 2-class male classification
4.5.3 Hybrid version experiment

In this experiment, we attempt to apply and evaluate the concept of the proposed hybrid version from Chapter 3 into the proposed method. We divide a face $X$ into 5 subsections, i.e., whole face, eyes, eyes and nose, nose and mouth, and mouth. Using Eq. 4.1 and Eq. 4.2, each subsections will have a class decision

$$\text{Class}(x) \in \{A, C, N\}$$

(4.5)

where $x \in \{\text{whole face, eyes, eyesnose, nosemouth, mouth}\}$ and with its respective $\delta_x$. The final class for the face is decided through majority voting among the class decisions of these subsections. Since there are 5 subsections and subsequently resulting in 5 class decision for each subsections, ambiguous cases, i.e., 2-2-1 vote for the 3 predefined classes might occur. In order to dissolve the situations when ambiguous cases occur, the $\delta$ value is utilized for the final class decision. We performed the following experiments for the hybrid version proposed method using different variations of solution during the ambiguous cases:

- During all cases, final class decision depends on the smallest averaged sigma value among all subsections (without majority voting).
- During ambiguous case, final class decision depends on the smallest averaged sigma value among all subsections (regardless of the number of votes).
- During ambiguous case, final class decision depends on the smallest averaged sigma value between the ambiguous classes that have equally maximum votes.
- During ambiguous case, final class decision depends on the summation of the averaged sigma value between the ambiguous classes that have equally maximum votes.

The boxplots for the 40-cross validation results for the original single tier proposed method and other variations of the hybrid version proposed method are shown in Figure 4.17. In the boxplots, the x-axis shows the experiments to be compared as listed in Table 4.4. While the results shown in Figure 4.17 vary depending on the participants, the boxplots showed that the performance of the proposed method in accuracy was not significantly improved with the hybrid version proposed method. Table 4.5 summarized the averaged improvement in accuracy between the original proposed method and different variations of the hybrid version proposed method. Among those variations of the hybrid version, there are only approximately 1% of improvement in accuracy.

The performance of the averaged accuracy of each subsection are shown in Figure 4.18. The figure showed boxplots of the averaged accuracy for each correctly classified subsections. In general, the whole face subsection from the hybrid version (i.e., the same as the original proposed method) yielded the highest accuracy. In the hybrid version, given a face $X$, if the class decision for the face is correct, the face subsection is also most likely be correctly classified. On the contrary, if the face subsection for face $X$ is wrongly classified, the other subsections that yielded lower accuracy compared to the face subsection did not contribute to change the wrong decision through majority voting. Therefore, the hybrid version did not improve in accuracy for this application compared with the original single tier propose method.
Figure 4.17 Results for hybrid version of the proposed method

<table>
<thead>
<tr>
<th>Color Code</th>
<th>Experiment</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>Proposed method (Single tier)</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>Hybrid proposed method using smallest sigma only (without majority voting)</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>Hybrid proposed method with majority voting and using smallest sigma only during ambiguous cases</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>Hybrid proposed method with majority voting and using the smallest sigma between ambiguous cases</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>Hybrid proposed method with majority voting and using the smallest summation of sigma values during ambiguous cases</td>
</tr>
</tbody>
</table>
Figure 4.18 Averaged percentage for the correctly classified subsections

Table 4.5 Averaged improvement in accuracy ($\bar{A}_q$) of the proposed method (Experiment 1) with different variations of the hybrid version proposed method (in percentage)

<table>
<thead>
<tr>
<th>Compared Experiment ($q$)</th>
<th>$\bar{A}_q$ (%)</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experiment 2: Sigma only (No majority voting)</td>
<td>0.8356</td>
<td>1.4874</td>
</tr>
<tr>
<td>Experiment 3: Smallest sigma during ambiguous cases</td>
<td>1.5403</td>
<td>2.1130</td>
</tr>
<tr>
<td>Experiment 4: Smallest sigma between ambiguous class only</td>
<td>1.0691</td>
<td>2.1350</td>
</tr>
<tr>
<td>Experiment 5: summation of sigma during ambiguous cases</td>
<td>0.6431</td>
<td>2.4210</td>
</tr>
</tbody>
</table>

4.5.4 Computational Order and Computational Time

The computational time of the proposed method and 2 other pixel image PCA-based conventional methods, i.e., Eisenthal pixel image method [19] and Turkmen method [20] are shown in Table 4.6.

In the training stage, the function of the highest computational order is the singular value decomposition (SVD) function that produced eigenfaces. As mentioned in Section ??, the computational order for SVD is $O(n^2m + n^3)$, where $n$ denotes the number of images and $m$ denotes the number of pixels in an image. Since $m \gg n$, the computational order can be rewritten as $m^2n$. Even though the computational complexity for all the compared algorithms are the same, the computational time for the proposed method was the shortest among all because the matrix in the proposed method involved is only 1/3 of the total size compared to the conventional methods, since the proposed method trained faces images from each class separately. Therefore, even
though in the proposed method the PCA step was repeated 3 times, i.e., for the attractive (A), common (C) and non attractive (N), during the training stage, the time it took was less when compared to the Eisenthal and Turkmen methods. Since, Eisenthal method involves an extra step, i.e., selecting the eigenfaces based on point-biserial correlation technique, therefore, the computational time for this method was the longest among all.

In the testing stage, the proposed method compared the reconstructed test image with the original test image to determine the class, the computation order for this Euclidean distance similarity measurement was $O(m)$. In the KNN classifier, the comparison of weights, $\Omega$, between the test image and the training images using Euclidean distance had the computational order of $O(n^2)$. For SVM classifier, the optimized computational order was $O(n)$. However, Eq. 2.4 was used for the image reconstruction stage in the proposed method and Eq. 2.5 was required to obtain the weights of the test images for the Eisenthal and Turkmen methods. Since the computational order for Eq. 2.4 and Eq. 2.5 were both $O(mn)$ and $m \gg n$, the computational order for proposed method and the conventional methods were $O(mn)$. The computational time for the proposed method was the longest among all since $m \gg n^2$ by comparing the computational order for classification step mentioned. The computational time for the SVM classifier was the shortest among all because the decision was achieved by comparing the distance from the 3 class support vectors.

Table 4.6 Computational Time

<table>
<thead>
<tr>
<th>Method</th>
<th>Computational Time</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Training Stage (s)</td>
</tr>
<tr>
<td>Proposed Method</td>
<td>1.2266</td>
</tr>
<tr>
<td>Eisenthal (Pixel) KNN</td>
<td>1.7794</td>
</tr>
<tr>
<td>Eisenthal (Pixel) SVM</td>
<td>1.8481</td>
</tr>
<tr>
<td>Turkmen KNN</td>
<td>1.5699</td>
</tr>
<tr>
<td>Turkmen SVM</td>
<td>1.6796</td>
</tr>
</tbody>
</table>

** Since the linear SVM scheme is used here, we are also able to transform the support vector input from the weight space ($\omega$ space) to the spatial space (image space) without Eq. 2.5 to achieve the optimized computational order $O(n)$.

4.6 Conclusions

A face classification method that utilized class-specific eigenfaces and image reconstruction method to classify both female and male faces based on personal attractive preference was proposed to examine the performance of the basic linear-based approaches in the abstract measurement face classification. The proposed method trained face images from each class separately to produce eigenfaces that contain more specific face information in each class than conventional PCA-based methods that trained face images from all the classes together. The proposed method classified the test image by locating the class that contain the most similar face features through the similarity between the test image and the reconstructed images using class-specific eigenfaces.

The results from the experiments varied depending on the participants, the proposed method outperformed all the other linear-based conventional methods for all the participants with con-
fidence level of 95% according to the Wilcoxon signed-rank test. In the 3-class classification, improvement in average accuracy ranged from 7.7% to 15.1% and 2.9% to 17.4% was achieved for female and male dataset, respectively. In the 2-class classification experiments, improvement in average accuracy ranged from 6.3% to 17.7% and 5.2% to 16.6% was achieved for female and male dataset, respectively. Furthermore, while the proposed method showed significantly improvement in accuracy, the computational order is the same as other conventional PCA-based pixel image methods.
Chapter 5

Conclusions and Future Works

Summaries of this research, conclusions reached, and suggestions for future researches are presented in this chapter.
5.1 Conclusions

Face classification approaches were designed to classify faces based on certain predefined criteria for various applications. In this thesis, we categorized the face classification approaches into 2 categories based on the criteria of the defined classes, i.e., exact measurement face classification that has crisp clear definition of face classes and abstract measurement face classification that has vague definition of face classes. Among the different face classification approaches, linear-based approaches are known for their high accuracy performance despite the simplistic nature. These attractive approaches have been applied into various face classification applications. The performance of the proposed linear-based approaches in both exact measurement and abstract measurement face classification is tested and examined.

For the exact measurement face classification, a hybrid fisher linear discriminant analysis (HFLD) methodology was proposed to overcome the limitation of the conventional FLD approach when handling faces of intruders in applications such as face recognition security system. The unique 5-tiers methodology that divides faces into 5 subsections, i.e., whole face, eyes, eyesnose, nosomeouth, and mouth, checks for consistency in the decision making stage to determine whether the testing face is from a user or an intruder.

In then experiments using the AT&T dataset, the proposed method surpassed the conventional FLD by 16.8% and 2.5% for face authentication and face identification, respectively. Also, when comparing with other hybrid methods, the proposed method achieved accuracy ranged between 4% to 10.4% and 1.2% to 5% improvement for face authentication and face identification, respectively. In the experiments using the MZ dataset to simulate an actual household security system, the proposed method achieved the overall recognition rate of 98.6% for face authentication and 96.3% for face identification.

For the abstract measurement face classification, a face classification based on personal attractive preference to classify both Asian female and male faces using class-specific eigenfaces reconstruction method was proposed. In the proposed methodology, each class was handled separately to obtain class-specific eigenfaces, which was different from the conventional PCA-based methods that train face images in all the classes as one huge training set. With the new proposed methodology, we were able to produce eigenfaces that contain more specific face information for each class when compared to the conventional methods. Then, the face classification was done by measuring the similarity between the test image and the reconstructed images using the class-specific eigenfaces.

From the experiments, while the accuracy results varied depending on the participants’ personal attractive preference, the proposed method outperformed conventional methods for all the participants, with a confidence level of 95% according to the Wilcoxon signed-rank test. In the 3-class classification, the proposed method achieved improvement in average accuracy ranges from 7.7% to 15.1% and 2.9% to 17.4% for the female and male dataset, respectively. Also, in the 2-class classification, the improvement in average accuracy ranged from 6.3% to 17.7% and 5.2% to 16.6% for the female and male dataset, respectively. The proposed method not only achieved significant improvement in accuracy, but its computational order remained the same when compared with other conventional methods.
5.2 Future Works

The simple yet high performance linear-based algorithms are worth further investigations. Applications on other image-based or even video-based classification to further challenge the basic linear-based approaches are interesting and possibly leads to the production of some practical and computational effective applications that is able to generate profit.

For the exact measurement face classification, the proposed HFLD face recognition methodology has been shown to yield satisfactory performance results. Research on the actual implementation of the proposed dual-function face recognition home security system may also be one of the directions for future exploration. The proposed HFLD can be implemented into an embedded system for further investigation. Implementing the algorithm using a low-level language such as C may reduce its training and recognition time.

For the abstract measurement face classification, there are still rooms for improvement on the accuracy rate for the proposed class-specific eigenfaces methodology. Other classifier methods for the methodology could be further explore. Also, using faces from other ethnicity for training and collecting personal attractive preference from participants of the same ethnicity for testing to further evaluate the proposed methodology may be one of the directions for future exploration.

For both the exact measurement and abstract measurement face classification, even though many linear-based approaches have been proposed for the past years, there are still possible alternations of the linear-based approaches that should be further examine in the future.
Bibliography


[63] V. Bruce, P. Hancock, and A. Burton, *In Face Recognition: From Theory to Applications*. Springer-Verlag, 1998.


Publications

Journal Papers


International Conference Papers


Domestic Conferences (in Japan)


Acknowledgements

First and foremost, I would like to thank Prof. Kiyoshi Tanaka, Prof. Hernan and Prof. Akimoto for all their hard works and efforts in guiding me through my studies at Shinshu University. Their guidance and invaluable comments have contributed significantly to this dissertation.

Naturally, this work has been influenced by many other researchers as well. I am in great debt to the anonymous reviewers for their invaluable comments and insights. My appreciation to Dr. Wong Kok Shiek at the University of Malaya who encouraged me to further my studies in Japan. I gratefully acknowledge the financial support provided by the Japanese Government Scholarship program - Monbukagakusho from April 2011 to September 2014. I thank my family and my husband for all the supports and loves they have provided especially during difficulties. Finally, I am grateful that I made many great friends through my study here and with their kind encouragement and supports that I achieved my dream to pursue my Phd in a foreign country.