KUMAMOTO UNIVERSITY

Pose Invariant Face Recognition Using Dominant Frequency Based Holistic Features and Statistical Classifier

by

I Gede Pasek Suta Wijaya

A thesis submitted in partial fulfilment for the degree of Doctor of Engineering

in the
COMPUTER SCIENCE AND ELECTRICAL ENGINEERING
Graduate School of Science and Technology

September 2010

Declaration of Authorship

I, I Gede Pasek Suta Wijaya, declare that this thesis titled, ”Pose Invariant Face Recognition Using Dominant Frequency Based Holistic Features and Statistical Classifier” and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed: ______________________________

Date: ______________________________

"The better you think, the better you will do;
the better you say the better you will hear; and
the better you do the better you get"

Taken from Trikaya Parisudha Concept:
three type of conducts that should be purified.

Abstract

Doctor of Engineering Thesis
In the Graduate School of Science and Technology KUMAMOTO UNIVERSITY
In Title of

Pose Invariant Face Recognition Using Dominant Frequency Based Holistic Features and Statistical Classifier

by I Gede Pasek Suta Wijaya

Face recognition is one of the most active research areas in pattern recognition, not only because the face is a human biometric system but also because there are many potential applications of the face recognition which range from human-computer interactions to authentication, security, and surveillance. The published methods of face recognition can be categorized in to three groups (Zhao et al., 2003). Firstly holistic matching method, which uses the whole face region as raw input to the recognition system; secondly features based (structural) matching methods which use the local features such as the eyes, noses, and mouth, and local statistics (geometrics and or appearance) as data input to the recognition system; and thirdly hybrid methods which use both whole face and local features as data input to the system. The PCA-based and LDA-based methods of face recognition that belong to the first category are well known and encouraging results have been achieved. However, both of them have their limitations: large computational costs, high memory space requirement, and retraining problems. In addition, the main disadvantage is the PCA projected features are lack of discrimination power because it removes the null spaces of data scatter that have the discriminant information.

This dissertation presents an approach to pose invariant human face image recognition. The proposed scheme is based on the analysis of discrete cosine transforms (DCT), discrete wavelet transforms (DWT), and moment analysis as global features extraction of face images. In detail, our proposed methods can be categorized as follows:

1. Dominant frequency features based face recognition. The scheme of this method is based on frequency analysis to obtain the features of face image

and multi-resolution metric to determine the similarity among the query face features and the training face features set. There are two main aims of the proposed method: to create holistic compact and meaningful features of face image without removing significant face image information, and to build a simple training algorithm that can solve the retraining problem of PCA-based face recognition. The proposed holistic features of face image are obtained by frequency analysis and quantization. The quantization will make the face features be not sensitive to small lighting variations.

2. Hybrid dominant frequency features based face recognition. Our previous work remains lack of discriminant power, which is shown by success rate still far from the maximum. It means the single frequency analysis is not adequately enough yet to represent the holistic information of any face pose variations. To address this problem, we propose an improvement of pose invariant human face recognition approach based on the analysis of the DCT and multi level Discrete Wavelet Transforms (DWT) of the face images. From both the DCT and DWT domain coefficients, which describe the face information, are fused as compact and meaningful features vector, using simple statistical measure and quantization called as the hybrid dominant frequency features. The main objective of the proposed method is to improve our previous method performance in term of recognition rate.

3. Statistical based face recognition. In order to improve the performance of our previous face recognition, we propose an features cluster which is derived from maximum a posteriori (MAP) discriminant called as modified LDA (MLDA). However, the propose features cluster does not works on quantized dominant frequency features because that features make the global covariance be singular. Therefore, we implements non quantized dominant frequency features. In this case, the MLDA is introduced to performed multi-pose face features cluster. There are several objectives of our proposed method. Firstly, to redefine our previous compact and meaningful face feature without removing significant face image information. Secondly, to build a simple classification technique that can classify face images to person's classes well. Thirdly, to make the MLDA-based training system that can solve the retraining problem. Finally, to know the effectiveness of DCT and wavelet analysis as face feature extraction when they are combined with the MLDA based classification. In last of this section, our proposed method

will work in color face image instead of grayscale in order to cover the skin information. The skin color face image is one of discriminant information which is available in the chrominance component and have to consider to get better performances. Consequently, it will take longer time processing than just working in grayscale.

4. Real-time face recognition using predictive LDA and alternative PCA. Previously, the global/holistic feature of face image, which is based on dominant frequency content, has been successfully implemented for face recognition. By using the global/holistic features concept as dimensional reduction of the face image could compress about 99.39% of the original size (i.e., less than 100 elements of 16384 elements) which gave good enough performance. It means the frequency analysis based global features is an efficient way for reducing the original data dimensional. Therefore, that holistic features concept, which is combined with the alternative PCA (APCA), predictive LDA (PDLDA), and integration of them, is implemented to create real-time face recognition. This proposed method is an alternative approach to face recognition algorithm that is based on global/holistic features of face image, which is combined with APCA and PDLDA to overcome large computational costs and retraining problem of the conventional PCA and LDA and with integration of APCA and PDLDA to improve features cluster. There are several objectives of the proposed method. Firstly, to prove that the global features of face image contains most of face classification information. Secondly, to define alternative PCA and redefine the predictive between class scatter (S_b) which have the same structure as their original but they have less computation complexity. Thirdly, to optimize the recognition performance using multi-stage classifier integration. Finally, to know the effectiveness of proposed methods compared with the best traditional subspace methods. In addition, the proposed method also work in color face image instead of grayscale to cover the skin information. Furthermore, shape face analysis performed by invariant moments is also included to get the holistic information of face pose variations.

In order to know the performance of our proposed methods, the experiments were carried out using several challenging face databases: the ORL database, YALE database, ITS-Lab. Kumamoto University database, INDIA database, and FERET database. Each database has special characteristics. From all of the

experimental results, the proposed methods have proved our hypothesis as described: firstly, the global features as a dimensional reduction of raw image has been successfully implemented with good achievement when Lq metric, MLDA, one-dimensional APCA, DLDA, and PDLDA as features clustering; secondly, PDLDA-based face recognition has been proved that they require less time processing for training and retraining; thirdly, the optimum recognition performance can be achieved by integrating the PDLDA and APCA or DLDA which the almost the same time processing; and the rest, the proposed method (APCA and PDLDA) outperforms over the recent sub-space methods (DLDA, 2DLDA, $(2D)^2$DLDA, 2DPCA, and $(2D)^2$2DPCA.) and the PDLDA can used as alternative solution to avoid recalculating the S_b and global mean of the retraining process of DLDA. In addition, the PDLDA is alternative solution for large size data clustering because it does not depend on the global means. Furthermore, the purposed method has low computational cost, requires little memory space, and can overcome retraining problem.

Keywords: Face recognition, frequency analysis, holistic features, multi-resolution metric, PCA, LDA, alternative PCA, and predictive LDA.

Acknowledgements

This work would not have been possible without supporting and encouragement of my supervisors, colleagues and lab mates. First of all, I would like to thank Prof. Keiichi Uchimura and Assoc. Prof. Zhencheng Hu who has given abundantly helpful and has assisted me in numerous ways including summarizing the contents, examining documents and in particular allowing me to joint and to publish some papers to the conferences and journal in order to get many inputs and corrections of my works.

Secondly, I would like to thank Japanese Government (Monbukagakusho) for the award of Monbukagakusho scholarship, which has supported me during my study and research in Kumamoto university, Electrical Engineering and Computer Science Department, Graduate School Science and Technology. I would also like to thank the Eelectrical Engineering Department, Faculty of Engineering, Mataram University in Indonesia which has permitted me to continue my study in Kumamoto University Japan.

Thirdly, I would like to thank the IEICE journal reviewer, IIEEJ journal reviewer, IJICIC journal reviewer, CSIE Conference reviewer, WIAMIS conference reviewer who has given some helpful comments and suggestions regarding to the presentation of the papers.

Fourthly, I would also like to thank my colleagues and lab mates especially to Frank Lamosa who has given much information regarding to recent technique that is published in IEEE journal and to his help for checking the English of my paper, to Wang for his discussion and information of the software and anything related to my work in ITS lab, and to Nakata Yuki for his supporting to provide some crop face database and helping to fulfil of many Japanese documents on applying Doctoral course in Kumamoto University.

Fifthly, I would like to gratefully acknowledge to owner of ITS. LAB., ORL, YALE, INDIA, and FERET face database and all people who have participated on building all face databases.

Finally, I cannot end without thanking my family, on whose constant encouragement and love, I have relied throughout my time at the University. It is to them that I dedicate this work.

Contents

List of Figures

List of Tables

Abbreviations

PCA	Principle Component Analysis
APCA	Alternative Principle Component Analysis
2DPCA	Two-Dimensional Principle Component Analysis
LDA	Linear Discriminant Analysis
DLDA	Direct Linear Discriminant Analysis
MLDA	Modified Linear Discriminant Analysis
PDLDA	Predictive Linear Discriminant Analysis
2DLDA	Two-Dimensional Linear Discriminant Analysis
2DLDA	Two-Dimensional Linear Discriminant Analysis
$(2D)^2$LDA	Two-Dimensional and Two-Directional LDA
$(2D)^2$LDA	Two-Dimensional and Two-Directional PCALDA
DCT	Discrete Cosine Transforms
DWT	Discrete Wavelet Transforms
HMM	Hidden Markov Model
CNN	Convolution Neural Network
MAP	Maximum A Posteriori
FAR	False Acceptance Rate
FRR	False Rejection Rate
EER	Equal Error Rate

For/Dedicated to/To:

Qinta my lovely daughter,
Mang Ayu my lovely wife, and
all of my lovely families.

Chapter 1

Introduction

Recently, face recognition is one of the most successful applications of image analysis and understanding. It has received significant attention during past few years even though very reliable methods of biometric personal identification system has been existed, e.g., fingerprints and retinal or iris based personal identification systems. This evidence is caused by two main reasons: the wide range of commercial and low-enforcement applications, and the availability of feasible technologies after 30 years of research, for instance, high speed and low cost of computer system with large RAM, free and open source library for computer vision (*openCV*), and sophisticated commercial tool. Table 1.1 presents list applications of face recognition in broadest area.

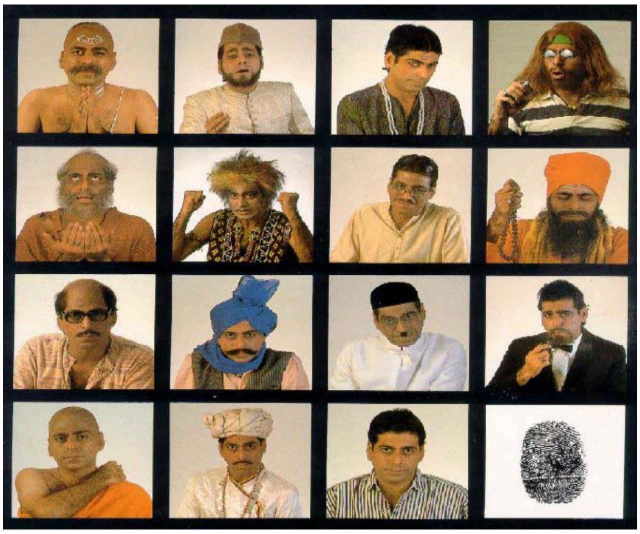
TABLE 1.1: Typical application of face processing including detection and tracking, recognition and verification, and personalized realistic [1].

No	Areas	Specifics Applications
1	Entertainments	Video games, virtual reality, training programs, human-robot interaction, human computer interaction, and family photo albums.
2	Smart Cards	Drivers licence, passports, voter registration, and welfare programs.
3	Information security	TV parental control, cellphone access, access control, password, PIN, internet access, and file encryption.
4	Low-enforcement and Surveillance	Advanced video surveillance, CCTV control, portal-control, post-event analysis, shoplifting, suspect tracking, and investigation.
5	Emerging	Tool for psychology study.

The idea of face recognition is inspired by ability of human being to recognize object or pattern based on training that has been performed continuously since childhood. Some researchers adopted that process to create any kinds of recognition systems, such as recognition system based on geometrical analysis, statistical analysis, neural network, etc. Face recognition is a matching process between a query faces features and target faces features. The process becomes difficult to be done because variations in a single face can be very large, while the variations between different faces can be quite small as shown in Fig. 1.1. Furthermore, face information depends on ethnicity and registration method (i.e., capture method, lighting condition, and device) as shown in Fig. 1.2. For example, each ethnicity has its own characteristics in term of skin, eye, nose, and face shape; the frontal face images contain clearer information than that of lateral face images; and the face images captured by digital provides more precise image than that of analogue camera.

From Fig. 1.1(a) shown clearly that the face variations of single person can be very large with putting different accessories. Therefore, the face can lie while fingerprint can not do at all. Regarding to pose and lighting variations as shown in Fig. 1.1(b), it also has large variations, but the human being still can understand that those are the face image of the same person.

The published methods of face recognition can be categorized in to three groups [2–4]. Firstly holistic matching method, which uses the whole face region as raw input of the recognition system; secondly features based (structural) matching methods which use the local features, such as the eyes, noses, and mouth, and local statistics (geometrics and or appearance) as data input of the recognition system; and finally hybrid methods which use both whole face and local features as data input to the system. The method that involves to the first category: principle component analysis (PCA), linear component analysis (LDA), probabilistic decisions based neural network, and their variations or combinations. The methods included in the second category are Dynamic link architecture, hidden Markov model (HMM), and convolution neural network (CNN). The methods included in the third groups are modular Eigenface, hybrid local feature method, flexible appearance model, and face regions and components. Among them, the PCA-based and LDA-based face recognition algorithms are well known and encouraging results have been achieved. However, both of them have their limitations: large computational costs, high memory space requirement, and retraining problems. Those systems



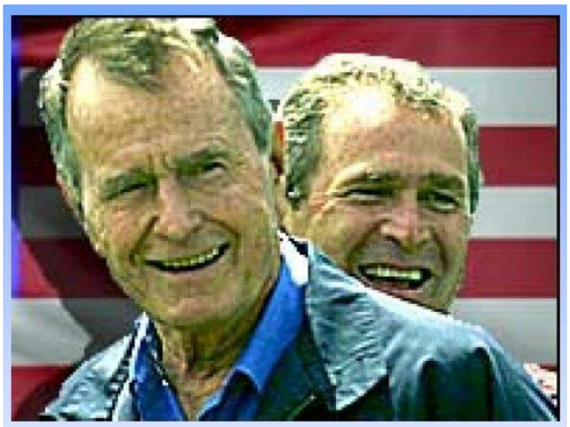
(a) Accessories variations

(b) Pose and lighting variations

FIGURE 1.1: Large variation in single person [1].

have to retrain all face image classes to get optimal projection when the new classes are added into the system. In addition, the main disadvantage is the PCA projected features are lack of discrimination power because it removes the null spaces of data scatter that still contains the discriminant information. The LDA works with assumption that the data samples have Gaussian distribution.

This thesis is organized as follows: chapter 1 describes about the introduction of this works including the previous works with their advantages and weaknesses, the brief theory of face recognition algorithm which involves the face features extraction and matching algorithm, and some proposed methods for solving weaknesses of the existed methods; chapter 2 describes dominant frequency features based face



(a) Father and Son [1]

(b) Among persons

FIGURE 1.2: Small variation Among persons.

recognition; chapter 3 describes the improvement of the chapter 2 in title of the hybrid dominant frequency features based face recognition; chapter 4 presents of the discriminant analysis called as modified LDA (MLDA) for improving the LDA based face recognition; chapter 5 describes the real-time face recognition using alternative PCA (APCA) and predictive LDA (PDLDA); and the rest concludes the work and proposes the future works.

1.1 Previous Works and Their Problems

The previous works, mostly related face recognition based on holistic or global approaches are presented in Refs. [5–9], as briefly described below.

Reference [5] proposed face recognition based on a combination of 2D-DCT analysis and face localization techniques for finding the global information (features)

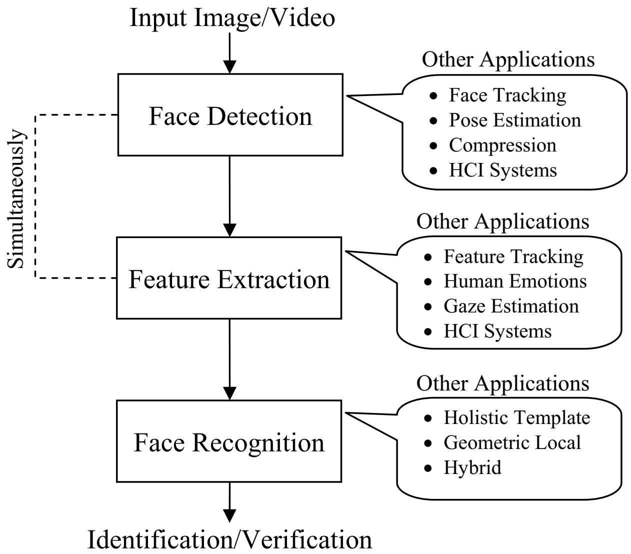


Figure 1.3: Configuration of generic face recognition/processing [10]

of face image, but it required eyes coordinate, which had to input manually for each face, to perform geometrical normalization. The global facial information was created by keeping small part of large magnitude DCT coefficients. Reference [6] also implemented 2D-DCT for face features extraction and selected the block size of 16x16, 32x32, and 64x64 of DCT coefficients as global features of face image. Then, the selected features were analysed by PCA to obtain the features cluster. In those approaches, the face information is based on the amplitude of DCT coefficients. However, the coefficients amplitude depends on pixels intensities of the image.

Reference [7] described face recognition based on wavelet packet tree analysis for frontal view of human faces under roughly constant illumination. The face feature was built by implementing wavelet packet tree analysis of the face bounding box and then calculating the mean and variance of sixteen wavelet coefficients matrices. The approach did not, however, work for frontal face view with little expression, accessories, and variant illuminations. In addition, it required constant illumination to make the face bounding box. Reference [11] implemented Daubechies wavelet analysis to get the global face information and classical PCA to classify the facial features to person's class. A mid-range frequency sub band image with resolution 16x16 was selected as face image the representation. However, the classical PCA need high computational cost and lack of power discriminations, as motioned previously.

Face recognition system based on frequency content that was extracted using discrete walsh transform (DWT) had been proposed in Ref. [9]. Face feature extraction was started from histogram equalization, next divided the face into some areas, then each area was normalized in the range [0-1], and finally two-dimensional DWT was implemented to get the features of face image. The matching process was performed using multilayer neural network. Pose invariant (face captured in range from 30 degree left to 30 degree right) face recognition system using neural network also was proposed in Ref. [12]. However, it took long time processing for training because neural network is difficult to tune the neurons for getting good enough classification engine.

Jacobs et al. [13] designed an image metric model for image retrieval based on Haar filter wavelet transform that implemented for image searching with scanned and painted image as the input data. The weakness of this method is in the data input. The scanned and painted image have large different in term of noise from face image which is usually taken by camera in the real situation. Kalman-face extraction application and tuneable parameter for face recognition was described in Ref. [14]. Each face was represented by a single feature vector that was extracted using Kalman face extraction. The similarity between query image and faces database were determined by the first order of Minkowski distance. The adaptive metric learning model that combines the generative and discriminative model has been developed successfully in Ref. [15].

Face recognition system based on Daubechies wavelets that had been implemented on FPGA as presented in Ref. [16]. Face feature was extracted by transforming each faces using 1-level wavelet transforms to generate average (A), detail vertical (V), detail horizontal (H), and detail diagonal (D) coefficients, then determining subspaces using eigenvector from those 4 subspaces, and finally computing their mean. Only Y of YUV colour component was processed. Face recognition method for frontal views of face under roughly constant illumination based on analysis of wavelet packet decomposition had been proposed. The face features was extracted by determining mean and variance of the decomposed wavelet coefficients. The matching process was done by probabilistic metric. Face recognition method based on colour information configuration which applied YIQ and YCbCr colour configuration (namely Y, I, Q, YI, YQ, IQ, YIQ, YCb, YCr and so on) was proposed in Ref. [17]. Face feature extraction was done using statistical technique that called as principle component analysis (PCA) and Fisher linear discriminant (FLD).

The most related works to our approach are the face recognition based on holistic or global matching methods as described in Refs. [2-19]. Ref. [2] presented a good review of face recognition method across pose, which explained that the PCA and LDA as well as their variations based approaches were widely used as face recognition because they had low computational complexity compared with the other methods. Ref. [3] described the evaluations of large-scale Chinese face database using several well know methods: PCA, PCA+LDA, Gabor-PCA+LDA, and Local Gabor Binary Pattern Histogram Sequence (LGBPHS). This database consisted of 99 594 images of 1040 subjects that had large diversity of the variations, including pose, expression, accessory, lighting, time, backgrounds, distance of subjects to camera, and their combined variations. The evaluation results showed that the algorithm based on LGBPHS outperformed over other three algorithms.

PCA [18–22] and LDA [23–26] are well-known scheme for feature extraction, dimensional reduction, and feature clustering. Those scheme and their variations have been successfully applied in face recognition. It was reported that the LDA was superior in face recognition to the PCA, because the LDA discriminated the data using both between class scatter and within class scatter of the training data. The Ref. [23] showed that the LDA provided higher power discriminant than that of PCA and the classification information of the PCA spread to all over to principle components while the LDA's concentrated in top few discriminant vectors. However, both PCA and LDA have their limitations: large computational cost, high memory space requirement, and retraining problems. Moreover, the PCA has poor discriminatory power while the LDA has singularity problems. Reference [27] proposed the combination D-LDA and F-LDA to cover the weakness of classical LDA term of poor discriminatory and singularity problems. However, it still needs high memory space and has be retrained when a new face class is registered into the system. Reference [28] implemented DCT to reduce data dimensional and only small number of the DCT coefficients were analysed by LDA. Reference[8] implemented optimization procedure and wavelet transforms to reduce the dimensional of face image and to employ a regulation scheme for the within-scatter matrix. It was reported that the Daubechies (Db6) was implemented to filter image to resolution 29 x 23. However, this resolution is still coarse and lack of frequency-resolution.

In term of data dimensional, the PCA and LDA can be classified in to one-dimensional (1D- PCA/LDA) [18, 23, 24, 27] which is vector-based analysis and two-dimensional (2D-PCA/LDA) [19, 22, 25] which is matrix-based analysis. The

2D-PCA outperforms over the 1D-PCA. However, the 2D-PCA requires more coefficients for image representation than 1D-PCA. The two-directional 2D-PCA [21], which works in both the row and column direction of image has been proposed in order to achieve higher and more stable accuracy as well as to decrease number of coefficients for image representation. However, they still have retraining problem.

In order to keep the two-dimensional structure of face image the 2D-PCA and 2D-LDA has been developed which the 2D-LDA gave better performance than that of the 2D-PCA. In order to get more compact features, the two-directional and two-dimensional PCA ($(2D)^2$PCA) [21] and PCALDA ($(2D)^2$PCALDA) [25] have been developed. The $(2D)^2$PCALDA out performs over the others algorithms. Therefore, we will compare our proposed methods to the $(2D)^2$PCALDA methods. The Ref. [29] studied extensively about the strength and the weakness of 1D-LDA and 2D-LDA. In our previous research, we also studied about the strength of 1D-PCA and 1D-LDA for face recognition with DCT based holistic features as raw input on the system. Moreover, we also developed an alternative 1D-PCA [30] that improved the PCA discrimination power and could overcome the PCA retraining problem. We also proposed weight 1D-LDA [31] that improved the class separable of the classical LDA. However, the LDA and its variations still have retraining problem because the between class scatter depends on the global mean, as described previously.

Mostly the previous works just worked in grayscale/intensity of image. They threw away the chrominance component where the face color information is available. In fact, the face image mostly covers with skin color, which is one of discriminant information of the face image. For example, if we want to match the black face image with the database which just contains white color face image such Chinese or Japanese, the system will can understand easily that the input image is not recognized because of the skin color. However, when the color skin information is not considered, the system will have possibility to recognize a face from white color face database. Therefore, we will consider the skin information for improving our previous works.

1.2 Brief Theory of Face Recognition System

Simply, face recognition is a matching process among a querying image and training images set that are stored in database. Li and Jain in Ref. [10] defines face recognition system as a process to identify faces present in images and video automatically. It can be done by either both of two methods: firstly, face verification (or authentication), which involves a one-to-one matching that compares a query face image against a template face image whose identity is being claimed; and secondly face recognition which involves one-to-few matching that compares a query image against to list of suspect of template face images.

A face recognition system generally consists of four main component: face detection, face enhancement, features extraction, training and matching as described detail below.

1. Face detection is defined as a computer technology that determines the locations and sizes of human faces in digital images. It is a segmentation to separate the human face from the background as illustrated in Fig. 1.4. In this case, the main aims to get accurately the face location that have to contain as many as human facial components, such as eyes, nose, mouth, and face outline.

(a) Input Image

(b) Detected Face

Figure 1.4: Illustration Face Detection.

2. Face preprocessing is regarding to image enhancement to get better image quality in term of it sharpness, clearness, and white balance (the balancing

of contrast and brightness). Other aim is to remove any lighting variations effect on face image capturing. It generally can be done by performing image stretching, histogram equalization, and geometrics normalization. The example of image preprocessing is shown in Fig. 1.5.

(a) Original Image (b) After hist. eq. (c) After stretching

FIGURE 1.5: Example of pre-processing of face recognition algorithms.

3. Features extraction is performed to provide effective information that is useful for distinguishing between face of different persons and stable respect to the geometrical and photometrical variations. The features of face image that is generally used as face features are : firstly, holistic/global information which is obtained by global transformation such as PCA, DCT, FFT, Wavelet, etc; secondly, the local features called as facial features which is obtained from analysis of facial components such as eyes, eye brow, nose, mouth and face outline.

4. Training/learning is process to tuning the classifier engine parameters which the main aims to get sophisticated engine that can determine the specific information of the input data, which can discriminate the subjects from difference class. For examples, the engine should give similar specific information of the subject of the same class and different specific information of the subject of difference class respectively. For the training, it is required set of data called as training data that have to inputed into the system for tuning the engine parameters. There are many types of training algorithms including: supervised learning which generates a function that maps inputs to desired outputs i.e. to approximate a function mapping a vector

into classes by looking at input-output examples of the function; unsupervised learning which models a set of inputs: like clustering; semi-supervised learning which is a combination of both labelled and unlabelled examples to generate an appropriate function or classifier.

5. The matching is a process to measure the similarity between the querying face to template faces that stored in database. The metrics is required in this process. Some common metrics that are generally used are L_1, L_2, Cosine, Correlation, and Mahalanobis distance.

1.3 Proposed Methods

In order to solve some problems that still existing in the previous proposed methods such as large computational cost, decrease memory space requirement and retraining problems, we propose some alternative methods as described below.

1. Dominant frequency features based face recognition [32, 33]. The scheme of this method is based on frequency analysis to obtain the features of face image and multi-resolution metric to determine the similarity among the query face features and the training face features set. There are two main aims of the proposed method: to create holistic compact and meaningful features of face image without removing significant face image information, and to build a simple training algorithm that can solve the retraining problem of PCA-based face recognition. The proposed holistic features of face image are obtained by frequency analysis and quantization. The quantization will make the features not be sensitive to small lighting variations.

2. Hybrid dominant frequency features based face recognition [34, 35]. The previous work still lack of discriminant power, which is shown by success rate still far from the maximum. It means the single frequency analysis is not adequately enough yet to represent the holistic information of any face pose variations. To address this problem, we propose an improvement of pose invariant human face recognition approach based on the analysis of the DCT and multi level Discrete Wavelet Transforms (DWT) of the face images. From both the DCT and DWT domain coefficients, which describe the face information, they are fused as compact and meaningful features vector, using

simple statistical measures and quantization called as the hybrid dominant frequency features. The main objective of the proposed method is to improve our previous method performance in term of recognition rate.

3. Statistical based face recognition [36–40]. In order to improve the performance of our previous face recognition, we propose an features cluster which is derived from maximum a posteriori (MAP) discriminant called as modified LDA (MLDA). However, the propose features cluster does not works on using quantized dominant frequency features because that features make the global covariance be singular. Therefore, we implements non quantized dominant frequency features. In this case, the MLDA is introduced to performed multi-pose face features cluster. There are several objectives of the our proposed method. Firstly, to redefine our previous compact and meaningful face feature without removing significant face image information. Secondly, to build a simple classification technique that can classify face images to person's classes well. Thirdly, to make the MLDA-based training system that can solve the retraining problem. Finally, to know the effectiveness of DCT and wavelet analysis as face feature extraction when they are combined with the MLDA based classification. In last section of this chapter, our proposed method will work in color face image instead of grayscale in order to cover the skin information. The skin color face image is one of discriminant information which is available in the chrominance component and have to consider to get better performances. Consequently, it will take longer time processing than just working in grayscale.

4. Real time face recognition using predictive LDA and alternative PCA [30, 31, 41–43]. Previously, the global/holistic features of face image which is based on dominant frequency content has been successfully implemented for face recognition. By using the global/holistic features concept as dimensional reduction of the face image could compress about 99.39% of the original size (i.e., less than 100 elements of 16384 elements) which gave good enough enough performances [33, 35, 38]. It means the frequency analysis based global features is an efficient way for reducing the original data dimensional. Therefore, that holistic features concept, which is combined with the alternative PCA (APCA), predictive LDA (PDLDA), and integration of them, is implemented to create real-time face recognition. This proposed method is

an alternative approach to face recognition algorithm that is based on global/holistic features of face image which is combined with APCA and PDLDA to overcome large computational costs and retraining problem of the conventional PCA and LDA and with integration of APCA and PDLDA to improve features cluster. There are several objectives of the proposed method. Firstly, to prove that the global features of face image contains most of face classification information. Secondly, to define alternative PCA and redefine the predictive between class scatter (S_b) which have the same structure as their original but they have less computation complexity. Thirdly, to optimize the recognition performance using multi-stage classifier integration. Finally, to know the effectiveness of proposed methods compared with the best traditional subspace methods. In addition, the proposed method also work in color face image instead of grayscale to cover the skin information. Furthermore, shape face analysis performed by invariant moments is also included to get the holistic information of face pose variations.

Chapter 2

Dominant Frequency Features Based Face Recognition

The face recognition is a matching process between query's face features and targets face features. The process becomes difficult to be done because the face features have large similarities in texture, color, shape, and pattern. The existed methods still have some difficulties in order to get the holistic features of face image. For instance, they require initial condition such as eyes, nose, and/or mouth position, and constant illumination which is not easy to complete. PCA and LDA are well-known methods for obtaining the holistic feature of face image. However, they require large time processing, large memory space, and singularity problem for small sample size. In addition, they have to retrain all of the data samples to get the optimum projection when the new data is added into the system.

In this chapter, an alternative approach to face recognition system that is based on frequency analysis and multiresolution metric is proposed to overcome large computational costs, high memory spaces requirement, and retraining problems of the eigenface-based method [6, 18]. There are two main aims of the proposed method: to create holistic compact and meaningful features of face image without removing significant face image information, and to build a simple training algorithm that can solve the retraining problem of PCA-based face recognition. The proposed holistic features of face image are obtained by frequency analysis and quantization. The quantization will make the face features not be sensitive to small lighting variations.

This chapter is organized as follows: section 2.1 describes the recognition algorithm of the proposed method, which consists of features extraction, training, and matching processes; section 2.2 presents the experimental results and their discussion; and the rest concludes the paper.

2.1 Proposed algorithm

The recognition algorithm can be shown in Fig. 2.1. It consists of pre-processing, features extraction, training, and querying processes. In the pre-processing process, grayscale transformation and histogram equalization are performed to get good quality face image. If the system receives a color image, it will automatically be converted to grayscale using the luminance model that is usually used by the NTSC (i.e. YIQ color space transformation). Then, each face image is normalized and equalized to remove non uniform-lighting effects during face capture. In the training process, the system creates target face features that represent the faces index and then save them into database as meaningful data. In this case, the face features are created based on low frequency components as described in the next subsection.

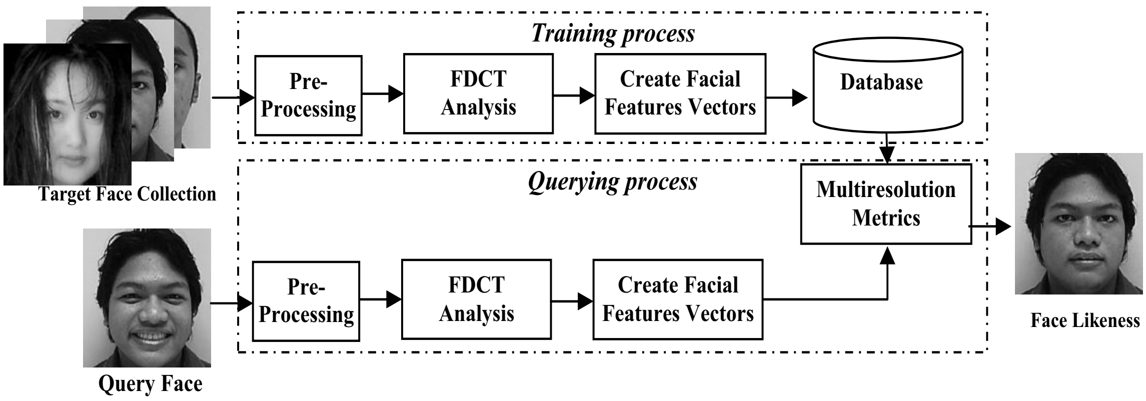


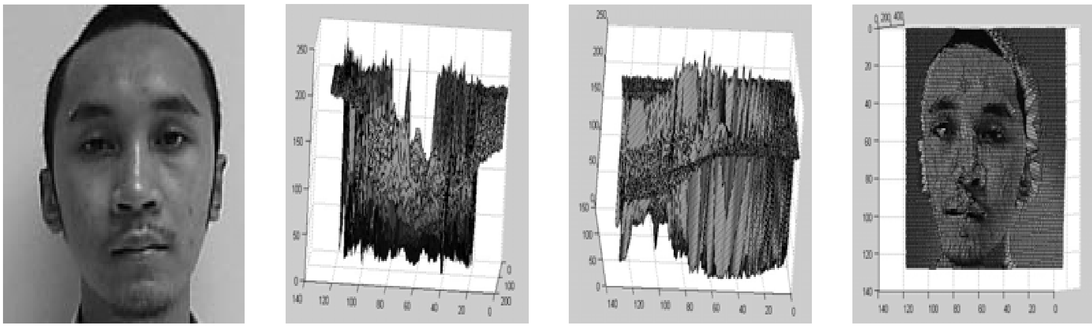
FIGURE 2.1: Face recognition diagram block.

In the querying process, the system calculates similarity score of a query's face features, which is obtained in the same way as in the training process, among the target's face features. It is computed using L_q metric that is described in the next subsection.

2.1.1 DCT Based Features Extraction

In this research, we develop a holistic approach for face features extraction based on frequency analysis of entire face. Our approach is different from the most closely related approaches as presented in Refs. [5–9, 13]. The Refs. [5, 6] implemented two-dimension blocked DCT on entire image as performed in JPEG compression and small number of the DCT coefficients were analysed by PCA and/or LDA. Ref. [13] implemented Haar basis function to create face features which is lack of smoothness and vanishing moments. Ref. [7, 8] implemented the wavelet- packet tree analysis for a bounding box face and then calculated the mean and variance of each band matrix coefficients. Ref. [9] proposed frequency content based features extraction which was started from histogram equalization, dividing the face into some areas, normalizing each area into range [0-1], and transforming them using two-dimensional discrete walsh transform to get the features of face image. Lately, multiple description image coding based on fractal was implemented for images transmission over unreliable channel [44]. It means that the multiple descriptions can be used as image features. However, our approach implements a 1-dimensional DCT analysis in the entire image without geometrical normalization and bounding process. The detail explanation of the 1D-DCT algorithms is described in Ref. [45].

This idea is inspired by human being ability that can readily and accurately extract patterns. For example, people who have studied A character can perfectly recognize each style of the A character, even though the A character is written in other shape, style, size, and color, such as a, A, **A**, *a*, etc. It means that human being has a sophisticated and complex technique to extract any patterns of A into the same information (features). In order to adopt this concept, the face image can be view as surface as shown in Fig. 2.2. It shows that the surface of face image consists of a collection of hills and valleys or a set one-dimensional signal. Theoretically, the frequency content of one-dimensional signal like human voice can be extracted accurately using several techniques such as FFT, wavelet, and DCT. Therefore, in this research, we implement one-dimensional DCT and simple statistical measures (mean and standard deviation) to extract the face image frequency. The DCT is chosen because it has good capability to bring the most important information of signal to low frequency components. Regarding to its strength, our hypothesis is an image features can be built using small percentage of DCT domain coefficients, which correspond to low frequency components and provide high correlated information of any face pose variations.



A V_{Row} V_{Col} A^{-1}

Figure 2.2: Image to vector transformation and plot to 3D

In order to generate frequency-matrix of the image in DCT domain, the 1D-DCT algorithms is implemented twice: on the rows and the columns of the matrix. In this case, we implement the 1D-DCT algorithms from Ref. [45] which is FFT-based DCT transformation as written below.

- Define *x(m)* vector from *y(m)* vector using the equation below

$$\begin{cases} x(m) = y(m) \\ x(N-1-m) = y(2m+1), \ (m = 0, ..., \frac{N}{2} - 1) \end{cases} \tag{2.1}$$

  Where N is vector y size

- Calculate the discrete Fourier transforms coefficients of *x(m)*:

$$X(n) = F[x(m)], \quad F \text{ is } \mathrm{F}FT \tag{2.2}$$

- Determine DCT coefficients using the equation below

$$Y(n) = Re\left[e^{-jn\pi/2N} X(n)\right] \tag{2.3}$$

The pseudo code of DCT decomposition for getting the frequency content of face image is written below:

func dctDecompose(*I*:array[0 .. *r*-1, 0 .. *c*-1])

```
    for row = 1 to r do
        F[row, 0 .. c-1]=1D-DCT(I[row, 0 .. c-1])
    endfor
    for col = 1 to c do
        F[0 .. r-1,col]=1D-DCT(F[0 .. r-1,col])
    endfor
    return F
endfunc
```

Where I is face image matrix and F is DCT domain of frequency matrix. The outcome of the DCT's pseudo code is plotted in Fig. 2.3.

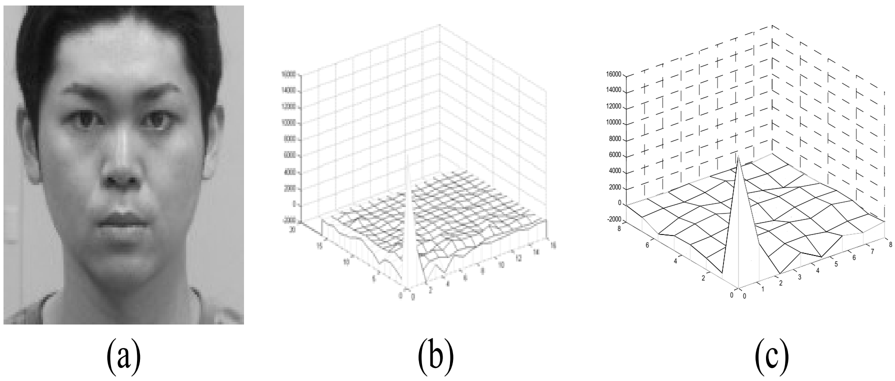


FIGURE 2.3: The output of DCT's decomposition pseudo code: (a) original image, (b) 16x16 selected coefficients, (c) 8x8 selected coefficients

As mentioned previously, each face image will be represented using a small number of the DCT coefficients, which are called as holistic features. The holistic features is a compact and meaningful features created by four steps: firstly, convert the frequency domain coefficients to vector using row ordering technique; secondly, sort the vector descending using quick sort algorithm, thirdly truncate a small number (m) of the sorted vector (i.e., $m \ll n$, n: size of original vector). Finally, the mean and standard deviation of selected coefficients are determined and then quantizing them to +1 (if the coefficients less-equal then 10^{-4}), -1 (if the coefficients more-equal then 10^{-4}), and 0 for otherwise. This quantization level will make quickly the matching process between query's and target's face features. The mean, standard deviation, and the quantized data are placed to a vector called holistic features, which is just called as a face features in the next presentation. The size of the face features will vary from 16 to 512 of 16384 coefficients. However, we will examine them to find the optimum of face features size. The features extraction process is performed on both training and query (probe) face images. However, in the training process, that is performed one time.

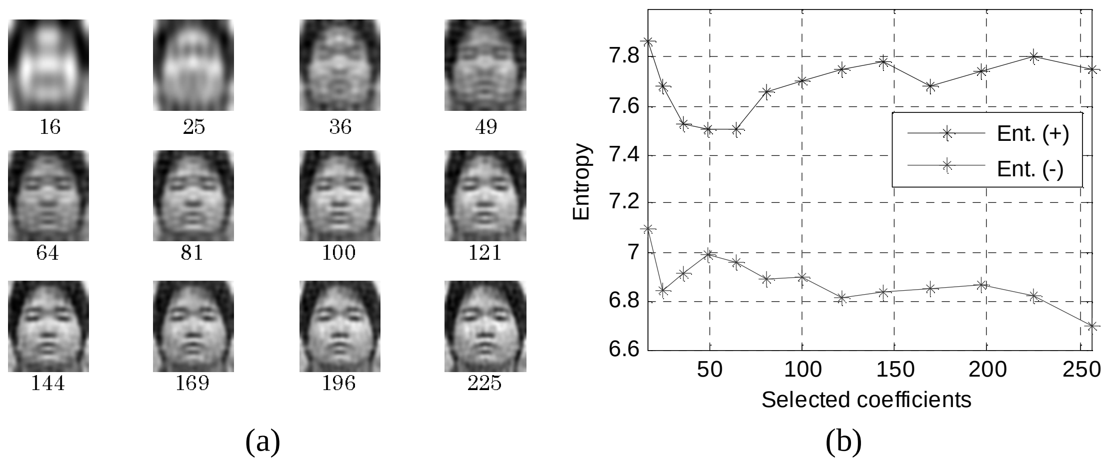


FIGURE 2.4: (a) The reconstructed images of the selected coefficients which the dimension ranges from 16 to 225 elements of the 16384 elements (original dimension), (b) the comparison of Ent (+) and Ent (-) as function of the selected coefficients dimension

Considering to the selected coefficients, if they are reconstructed into the face images, the reconstructed face images will be different. However, we can still understand that they are the face images, as shown in Fig. 2.4(a). In addition, the Fig. 2.4(b) shows that the entropy positive (Ent (+)) values are higher than (Ent (-)) values, where the Ent (+) is the average entropy of reconstructed image set when only the most information is considered, see Fig. 2.4(a). Meanwhile, the Ent (-) is the entropy of reconstructed image when the most information is removed. This illustration proves that the most information of image exists in low frequency components and it shows that the DCT domain coefficients contain powerful information.

2.1.2 Training Process

As shown in Fig. 2.1, the training process consists of preprocessing, DCT analysis, and features extraction. However, in the training process just the target faces features are created and save them into database as meaningful data. In this case, just the frontal face image is considered as training face because it contains clearer information than that of others do. We suppose that one face training per class is enough for representing the face class because the compact face features consists of dominant frequency components of the face image. From the frequency analysis (i.e. the DCT), we can get the collocation close to identical face features for any limited pose variations of single face. It can be proved by comparing correlation coefficients of both compact face features and original images of any face variations in a single face. Suppose that the Q1 face image of Fig. 2.6(b) as the training face and its remaining as the query faces. Next, the correlations coefficients between Q1's face features and the remaining face features (Q2's ... Q10's) are determined and we get almost the same correlation coefficients values. However, the correlation coefficients of the original faces have large different values for all face variations, as shown in Fig. 2.5.

This training process can solve the PCA's retraining problem because it does not required the global mean of training samples for creating the face features as performed on PCA training process [18, 28]. For instance, if a new class comes to the system, the proposed training system will train just its class without considering the previous class. However, in the PCA system we have to consider the previous classes in order to get the global mean and optimum projection matrix.

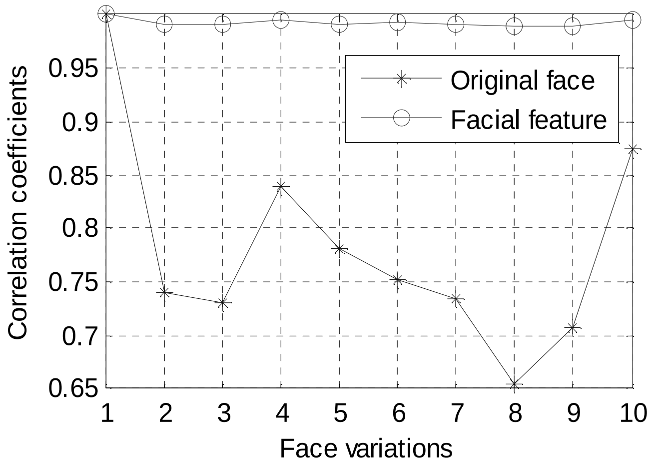


Figure 2.5: The correlation coefficients of the original image and face features for ten face variations of ITS-Lab. face database

2.1.3 Matching Process

There are many metrics methods that can be used as matching process such as L_1, L_2, L_q, probabilistic, Euclidean distance, adaptive metric learning, and Mahalanobis distance. In this case, we implement a combination of L_2 and L_q metric.

Fast L_q metrics equation was introduced in the Ref. [13] as shown below:

$$\|Q,T\|_q = w_{0,0}\{Q[0,0] - T[0,0]\} - \sum_{i,j:\overline{Q}[i,j]\neq 0} w_{i,j}\{\overline{Q}[i,j] = \overline{T}[i,j]\} \quad (2.4)$$

Where $\overline{Q}[i,j]$ and $\overline{T}[i,j]$ are the query and target face features. The term $\overline{Q}[i,j] = \overline{T}[i,j]$ will return 1 if $\overline{Q}[i,j]$ is equal with $\overline{T}[i,j]$, and 0 if otherwise. This comparison make quickly the calculation process, because the computer system performs faster the comparison than arithmetic operations.

The face features vectors consist of statistic features and frequency features, therefore a modification of Eq. (2.4) is applied for calculating similarity level as written in Eq. (2.5).

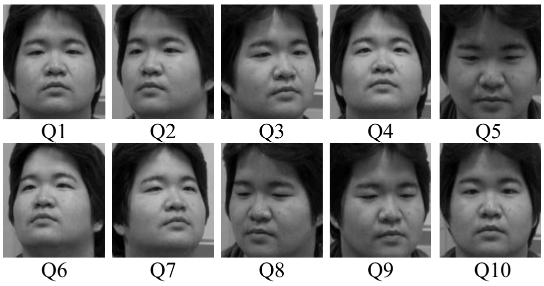
$$\|Q,T\|q = \sqrt{\frac{1}{2}\sum_{i=1}^{2}(v_i^q - v_i^t)} - \sum_{j:\overline{Q}[j]\neq 0}\{\overline{Q}[j] = \overline{T}[j]\} \quad (2.5)$$

Where v is statistic face features, Q is query frequency features, and T is target frequency features. The smallest score is concluded as the matching criteria (the best likeness).

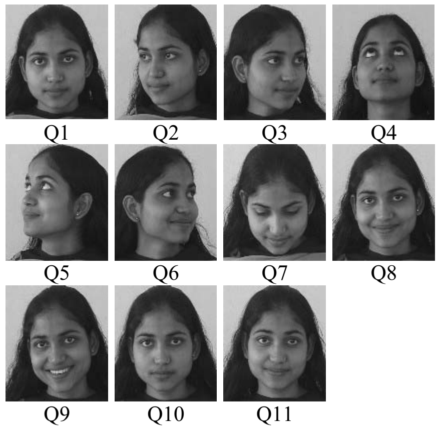
2.2 Experiments and Results

The experiments were carried out with four face databases: the ITS-Lab. Kumamoto University database, the EE-UNRAM database, the Indian database [53], the ORL database [46, 47]. Each database has special characteristics, which are described below. Totally, all databases consist of 204 classes and 2036 face images. The ITS-Lab. database consists of 48 Japanese people and each person has 10 pose orientations, as shown on Fig. 2.6(a). Totally, the face images are 480 samples. The face images were taken by Konica Minolta camera series VIVID 900 under varying light conditions.

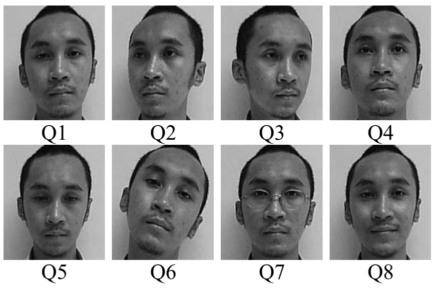
The Indian database consists of 61 people (22 women and 39 men) with each person having eleven pose orientations as shown on Fig. 2.6(b): looking front, looking left, looking right, looking up, looking up towards left, looking up towards right,



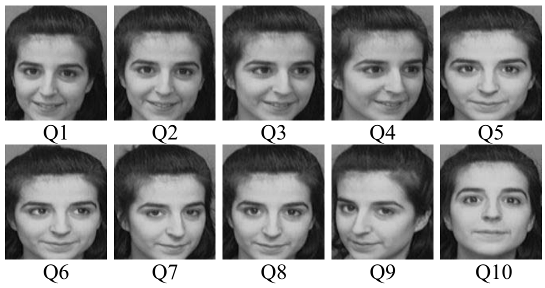
(a) ITS. Lab. database



(b) INDIA



(v) UNRAM



(d) ORL

Figure 2.6: Example of face pose variations.

and looking down. The Indian database also included some emotions: neutral, smile, laughter, sad/disgust. Totally, the face images are 671 samples.

The EE-UNRAM database was created by computer informatics Laboratory, Department of Electrical Engineering Mataram University which consists of 40 Indonesian people and each person has 8 pose orientations, as shown on Fig. 2.6(c): looking front, looking left about 300, looking right about 300, looking up, looking down, and wearing accessories such as glasses. Totally, the face images are 320 samples.

The ORL database was taken at different times, under varying lighting conditions with different face expressions (open/closed eyes, smiling/not smiling) and face details (glasses/no glasses), as shown on Fig. 2.6(d). All of the images were taken against a dark homogeneous background. The Faces of the subjects are in an upright, frontal position (with tolerance for some side movement). The ORL database is a grayscale face database that consists of 40 people, mainly male. Totally, the face images are 400 samples.

All the experiments were performed in the mentioned databases with the following assumptions:

1. The face image size is 128x128 pixels (28 pixels/cm) with representing using 24 bit for color image and 8 bit of for grayscale image per pixel.
2. The frontal face image is locked as a training face, for example: Q1 is selected for ITS-Lab., INDIA, UNRAM, and YALE face database, and Q5 is selected for ORL.
3. The training face is included in the query process because there is possibility that the query faces are almost or exactly the same as the training face in the real time system.
4. All face images in the same database must have the same background.

The first experiment investigated the effect of features vector size to success rate. This test was performed in the ITS-Lab. database, DCT-based features extraction, and one training faces per class. The success rate slightly decreases when the size of face features increases, as shown in Fig. 2.7. The face features size that is sufficient to achieve a good success rate ranges from 64 to 128 elements. This result can be achieved because the most face information exists in low frequency components (as proved in Fig. 2.4.). Therefore, the face can be distinguished by just considering them. In addition, face images of ITS-Lab were captured by good camera and the face pose variations are not very large, as shown on Fig. 2.6(a). In other side, the proposed system requires a short training and querying times by about 7.109 second and 0.682 millisecond while the PCA needs by about 18.36 and 0.061 second respectively, when the test was done in personal computer with specification: Intel Pentium M processor 1.7 GHz, 400 MHz FSB, and 768 MB of RAM.

The second experiment was performed to evaluate the stability of the proposed method against face variations. It was performed in the India database, one training faces per class, DCT-based face features, and face features of 64 elements. The result shows that our method gives better and more stable success rate than that of classical PCA, as plotted in Fig. 2.8(a). This result can be achieved because DCT has good energy compactness; therefore an image features can be represented by little mid-range frequency components. Furthermore, the DCT based face features of any face pose variations in a single face are almost identical (see Fig. 2.5).

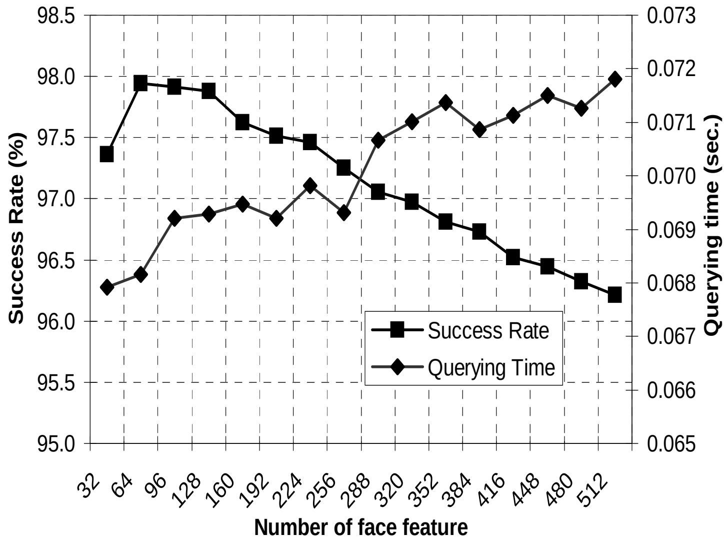
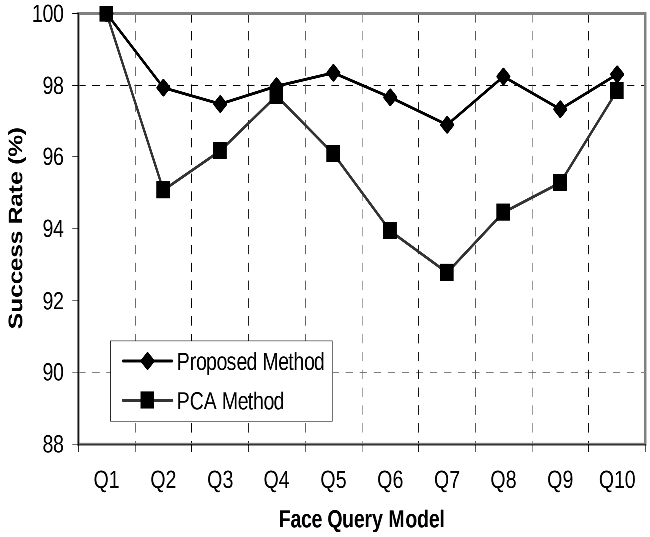


FIGURE 2.7: Success rate and time querying vs. face features size

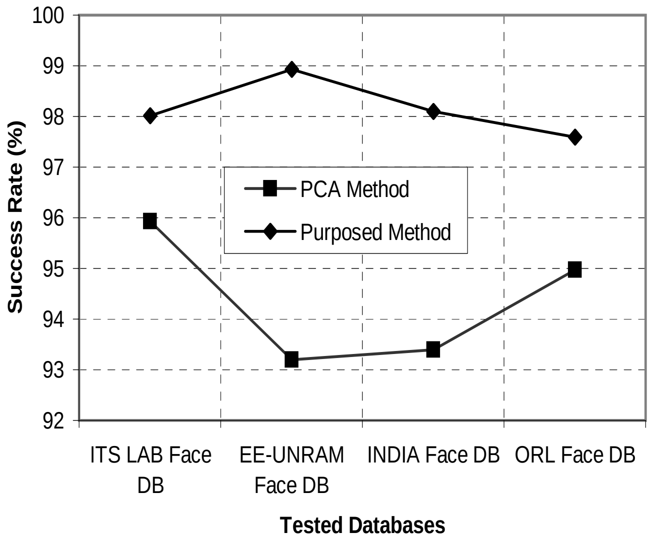
The next experiment is carried out to investigate the robustness of the proposed method, which is performed in the data from four face databases, one training faces per class, and face features of 64 elements. The difference of success rate of the purposed method for all tested face databases is not more than 1%, as shown in Fig. 2.8(b). It means that the proposed method is robust for the tested face databases. In addition, this result also shows that the proposed method is superior to classical PCA.

The last test was performed to know the accuracy of the proposed method compared to the classical PCA method. In this test, we investigated two important parameters, namely false rejection rate (FRR) and false acceptance rate (FAR). The FAR is the probability of unauthorized user to be falsely as accepted or recognized as legally user. The FRR is the probability legally registered user to be falsely rejected by the system. If the value of the FAR and the FRR is equal, this point is called as equal error rate (EER). The System, which performs perfectly classification, is denoted by 100% true positive rate and 0% false positive rate or the value of EER is very small or close to zero.

The experimental result is plotted using receiver operating characteristics (ROC) as shown in Fig.2.9. The results show that the proposed method (DCT) performs better than classical DCT+PCA for tested face databases, which is denoted by the smaller the equal error rate (ERR). The proposed method is not compared with LDA methods because LDA-based face recognition must be trained with more



(a) (a)stability success rate to face poses of INDIA face database



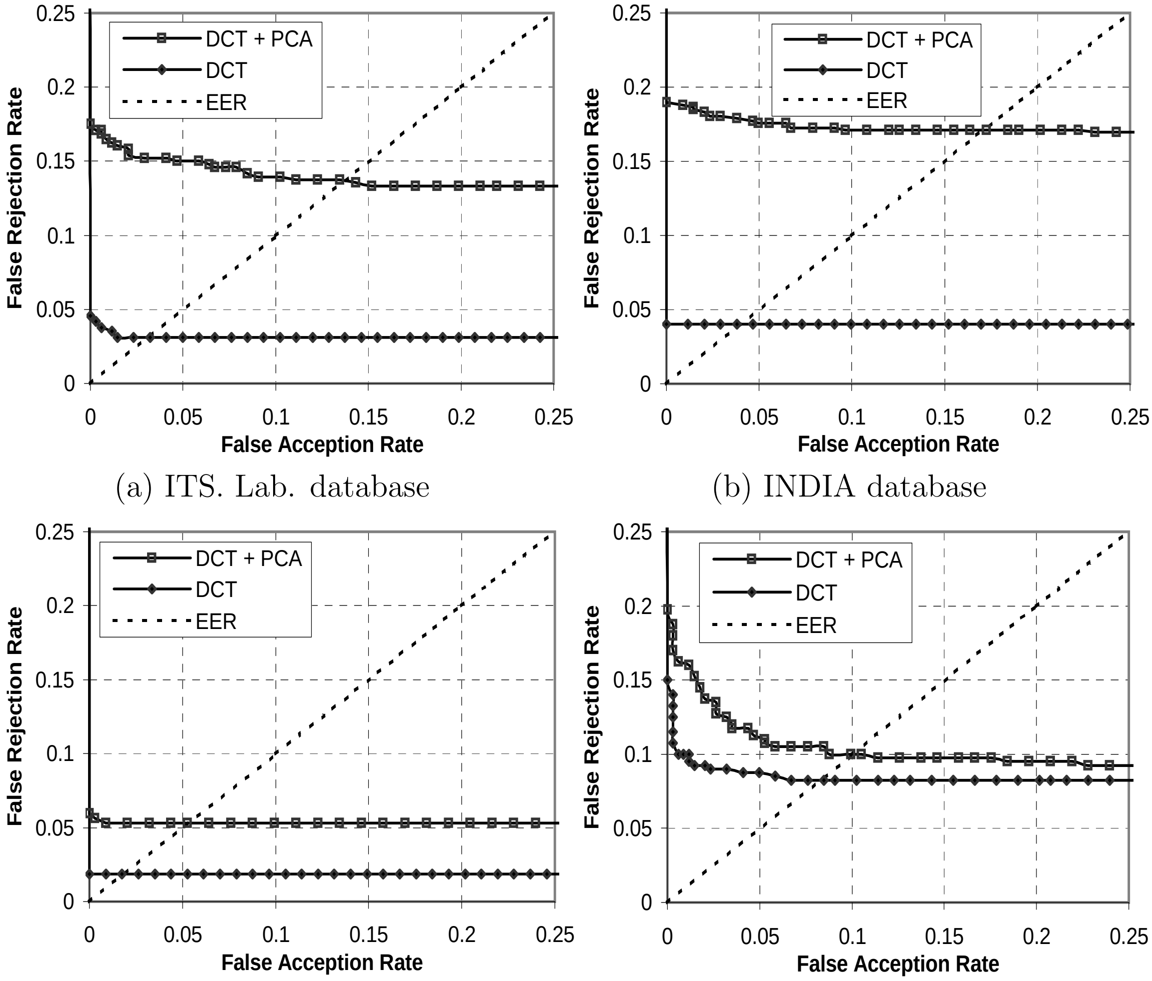
(b) (b) robust success rate in four face databases

FIGURE 2.8: The comparison of success rate.

than one training face image per class [23, 28]. Meanwhile, all of the experiments are performed using one training face image per class.

2.3 Discussion

The proposed method has been implemented successfully and given good performance, robustness, and stability without requiring geometrical normalization. This performance can be achieved because some reasons:



(a) ITS. Lab. database

(b) INDIA database

(c) ORL. database

(d) FERET database

FIGURE 2.9: ROC curve of proposed method compared to the other systems.

1. The most signal information exists in low-frequency components, while the high-frequency components impart the flavour or nuance of signal. Consider the human voice, if we remove the high-frequency components, the voice sounds different, but we can still tell what is being said. However, if we remove enough of the low-frequency components, we will hear gibberish. Based on this perception, therefore, the face image information can be represented by the low-frequency components.

2. The compact face features concept consists of the quantized coefficients of the low frequency components, which represents the dominant frequency existence of signal. Moreover, the frequency existence does not depend on the pixels intensities but the shape of data.

3. DCT has good capability in separating the frequency components of the signal. By the DCT, the most signal's information is brought to the low frequency components, which exists in the top left of transformed coefficients.

4. The DCT coefficients are uncorrelated with their counterpart frequency indices. It means that if most of non-dominant frequency component is removed it will not significantly decrease the quality of its information.

5. Fast DCT algorithm needs $N\, log2\ N$ time complexity, where N is the number of data.

The proposed methods can create compactly face features which needs small memory space (94 bits). In this case, the format of face features consist of statistical features that are saved using double data type and quantized data that are saved using bit data type. It can be done because the quantized data only have two combinations: +1 and -1. Therefore, each face image can be represented using 0.39% of all face frequency coefficients. The PCA analysis needs more training time because it makes the face features using eigen vector analysis of covariance matrix of face training set. Moreover, the DCT based features that implemented in the PCA must be allocated using double memory data type (2 byte) because the features do not contain quantized values. It means that our hypothesis is proven by these experimental results.

2.4 Conclusions and Future Works

The proposed method has been successfully implemented and has given good performance, which the optimum success rate about 98% and small EER. This system also can reduce training time about 61.28% of PCA training time, and need little time querying by about 0.68 millisecond. The significant outcome is that, this system requires small face features about 5.87% of PCA face features, which give bigger success rate than PCA. This process needs some improvements such as applying cluster face analysis to make face group based on skin color detection and implementing the moment to detect any angle of face capturing.

The next research will focus on clustering analysis and applying hybrid frequency analysis (DCT and DWT) in creating the face features for improving the performances. Moreover, the proposed method will be done in more face databases such as Yale database in order to know its accuracy and reliability.

Acknowledgement

The authors gratefully acknowledge to Kumamoto university which supported our publication of this paper, to the owner of ITS-Lab., ORL, and EE-UNRAM face databases. Further more, the gratefully acknowledge to IJICIC reviewers who gave some helpful comments and suggestions for improving the presentation of this paper.

Chapter 3

Hybrid Dominant Frequency Features Based Face Recognition

Our previous work remains lack of discriminant power, which is shown by success rate still far from the maximum. It means the single frequency analysis is not enough yet to get the holistic information of any face pose variations. In other word, the DCT still have few missing dominant information of face image. However, our previous method has better performance than that of Eigenface (PCA) for both success rate and time processing. Furthermore the proposed method does not require the retraining problem because features extraction does not depend on the global mean as done by the PCA algorithm.

To address this problem, we propose an improvement of pose invariant human face recognition approach based on the analysis of the DCT and multi level Discrete Wavelet Transforms (DWT) of the face images. From both the DCT and DWT domain coefficients, which describe the face information, we fuse them as compact and meaningful features vector, using simple statistical measures and quantization. This feature vector is called as the hybrid dominant frequency features. The main objective of the proposed method is to improve our previous method performance in term of recognition rate. These goals can be realized by implementing the concept of hybrid dominant frequency features that is obtained by fusing the low-frequency components of both the DCT and wavelet domain coefficients. The hybrid dominant frequency features concept is a compact and meaningful global representation of face information. We choose the low-frequency part because theoretically most signal information is available in the low-frequency components.

In the DCT domain, the low-frequency components exist in the low coefficient indices (i.e., top left region) and in the DWT domain, the low-frequency components exist in the approximation coefficients. Therefore, with this feature, the proposed method only needs a short time for the training and querying process.

This chapter is organized as follows: section 3.1 explains the proposed face recognition algorithms which involve the face feature extraction, training, and matching algorithms, section 3.2 presents the experimental results and compares our results to results of the Eigenface method, section 3.3 describes the results discussion, and the rest concludes the paper.

3.1 Proposed Algorithm

The face recognition algorithm discussed in this chapter is the same as shown in Fig. 2.1. However, in this chapter, two frequency analysis (DCT and DWT) is implemented which performs simultaneously to obtain holistic features of face image, as shown in Fig. 3.1. As done previously, the proposed algorithm works on both color and grayscale image. If the system receives a color face while the face database is in the grayscale format, it will automatically convert the color face image into the grayscale format using a luminance model, which is used by NTSC. Next, the image is equalized to remove non uniform-lighting effects during face image capture. The training process builds target face image features, which represent each face index and then save them into a database.

In the recognition process, query face features is built in the same way as in the training process but it is not saved. Next, the combination of L_2 and L_q metrics is used to calculate the similarity between the query feature and target feature sets. The smallest score is selected as the best likeness.

3.1.1 Hybrid Feature Extraction

The hybrid features of face image is a holistic face feature which is determined using both of DCT and DWT analysis of entire face image. Our approach is different from the most closely related approached presented in Ref. [5–8]. Our approach builds the holistic compact and meaningful face feature using simple statistical

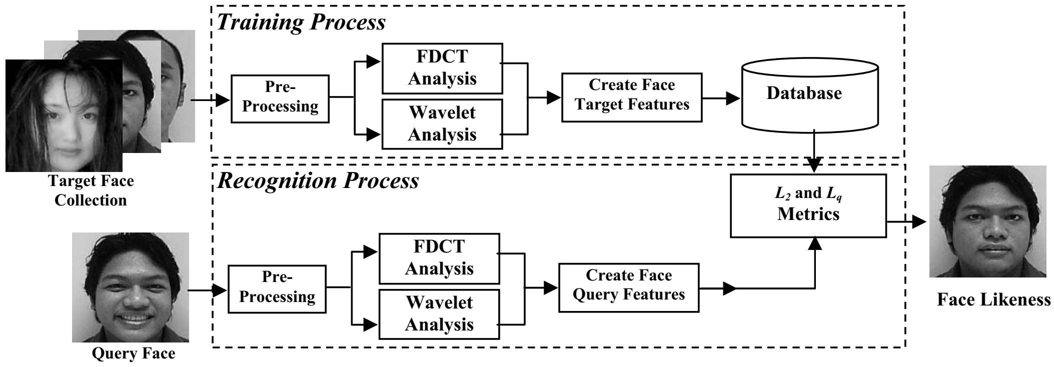


FIGURE 3.1: Proposed face recognition diagram block

measures and quantization of dominant frequency content of face image. It means that the face feature does not contain amplitude value of coefficients but the sign of frequency existence in entire face image. We do not use the amplitude as face feature because the amplitude value depends on pixel intensities of the image which, in turn, depends on lighting conditions, whereas the frequency depends on shape/texture of image.

Hybrid dominant frequency features extraction consists of three main processes: pre-processing, frequency analysis, defining features as described below:

1. Pre-processing.

   To remove non uniform lighting effects during face capture, we implement standard histogram equalization and normalization. The standard/common histogram equalization can be found in Ref. [48]. Because the face image is represent using 8 bits (0-255 pixel intensity), we normalize the face image into range [0-1].

2. Frequency analysis.

   In this step, the equalized and normalized face image is analysed by 1D-Fast-DCT and 2D-DWT for constructing two face frequency matrices. First, the DCT domain frequency coefficients is determined using the 1D-FFT based DCT algorithm which its decomposition procedure and pseudo code have been described in the Chapter 2 subsection 2.1.1.

   Next, multi-level DWT is implemented to decomposes the face image to wavelet domain frequency coefficients. The multi-level DWT analysis can be performed by repeatedly implementing the classical DWT, as shown in the following pseudo code.

func dwtMDecompose(I:array[0..r-1, 0..c-1])
  for res = 1 **to** n **do**
    $[A,H,V,D]_{res}$=dwt *(I,h,g)*
    $I = A_{res}$;
  endfor
  return *[A, H, V, D]*
endfunc

The classical DWT algorithm of the image is given by:

$$\begin{aligned} A &= [h * [h * f]_x \downarrow_2]_y \\ H &= [g * [h * f]_x \downarrow_2]_y \\ V &= [h * [g * f]_x \downarrow_2]_y \\ D &= [g * [g * f]_x \downarrow_2]_y \end{aligned} \tag{3.1}$$

Where * denotes convolution, $\downarrow_2$ represent down sampling in the x and y direction, g and h are high-pass and low-pass filter respectively. A, H, V, and D are the approximation, horizontal, vertical, and diagonal wavelet coefficients respectively. The data structure of the wavelet domain coefficients, which is the out come of above function, is [A_n,H_n,V_n,D_n,H_{n-1},V_{n-1},D_{n-1},H_{n-2},V_{n-2},D_{n-2}, ... , H_1,V_1,D_1], where n is the decomposition level.

3. Define the face features

A face image feature is built by selecting a small number (m) of both the DCT and wavelet domain coefficients (i.e., ($m/2$) form the DCT and ($m/2$) form the wavelet), which have large magnitude. Next, from the selected coefficients, we determine the means and standard deviations and then quantize the selected coefficients to +1, 0, or -1. The quantized value represents frequency dominant existence of face image that is indicated by the magnitude values of the selected coefficient is more than zero. Therefore, the quantization is performed by the following equation:

$$f_i = \begin{cases} +1, & if\ c_j > 10^{-4} \\ -1, & if\ c_j < 10^{-4} \\ 0, & otherwise \end{cases} \tag{3.2}$$

where j=1, 2, 3, ..., m. The 10^{-4} is chosen as threshold value because it is usually used as convergence criteria in the numerical analysis, such as obtaining root of polynomial equation using Newton Rapshon algorithm. Then the means, standard deviations, and quantized data is placed into a vector, $F_i = [\mu_i^{DCT}, \mu_i^{DWT}, \sigma_i^{DCT}, \sigma_i^{DWT}, f_i^{DCT}, f_i^{DWT}]$, where i is i-th class of face. In this case, the dimension of the feature, F, is m+4 and the f_i's elements consist of quantized value and its index that are saved using link-list data structure (*struct* in C++ programming). This face image feature has two advantages: it requires little memory space and needs short time in the matching process.

The proposed feature extraction process has some advantages in the training of face images, such as:

- It is simple because it does not need to determine the eigenvalues and eigenvectors of covariance matrix, as performed in the PCA and LDA.
- If a new class member is added to the set, the system just calculates its features, and then adds it to the vector class set, while PCA have to recalculate covariance of all of the classes and re-perform the eigen analysis of the newest covariance to obtain optimum matrix projection. These recomputing process make the PCA system require large computational cost for retraining.
- Consequently, the computation complexity of our proposed features extraction is less than that of the PCA and LDA.

Regarding the India face database example (Fig 2.6(b)), it has large variations in single face, which is showed by large differences correlation coefficients (see Fig 3.2, on the original part). These variations are caused by face angles. However, human being can identify that they have the same feature. The proposed feature extraction can be used to get their similar feature because the features consist of the dominant frequency components. Furthermore, the dominant frequency components do not depend on pixels position (face angle) but the pixels sequences (wave shape). As an illustration, consider the one-dimensional discrete signal with the frequency dominant ranges from 50 to 180 Hz. If the signal is shifted and its amplitude is gained, these processes will not affect the frequency content. This perception is similar to the two-dimensional signal (image). Even though the pose variations in single face are large, they still contain the same dominant frequency contents. It can be proved by calculating correlation coefficients of the original image and the face features. For example, given the frontal face (Q1 of Fig.2.1(b)) as a training and the remaining as querying (Q2, ..., Q11), if correlation coefficients of training face among the querying faces for both the original and face features are computed, we will get the result as shown in Fig. 3.2. This results show that the correlations of the face features are almost the same for all face poses variations even though their hair and background are included in the calculation. However, the correlations of the original faces have large differences for all face poses.

Based on this perception, we suppose that the proposed method can be used as a basis of pose invariant face recognition because the identical face features of any face pose variations in a single face can be obtained using the proposed feature extraction. However, in this case, the pose invariant is defined as the face view variations, which their face angle are restricted from -30 to 30 degree for up-down and left-right respectively and also have small visual expression. By this angle constraint, the two eyes still exist in the image.

Consider to the face features correlation coefficients as shown in Fig. 3.2, they show highly correlated values between the training and the querying feature without geometrical normalized and localization. Therefore, based on this reality, the face recognition can be performed using one face training per class. However, the frontal face must be locked as a training face because it contains complete face elements, such as eyes, mouth, ears, and nose. Moreover, human being can recognize more easily who a person is, using the frontal than using lateral face image. We also suppose that the more training faces are given the better recognition rate is gotten. However, they will require more memory spaces for saving them.

3.1.2 Matching Method

The main goal of this process is to calculate the similarity level between the query feature and target feature set. There are many metrics that can be used to calculate the similarity. Some of them are L_1, L_2, L_q, probabilistic, Euclidean distance, adaptive metric learning and Mahalanobis distance. In our system, the implemented metric is a combination of L_2, and L_q (multiresolusion metric) metrics. A detailed explanation of the fast multiresolusion metric was described in Ref. [13].

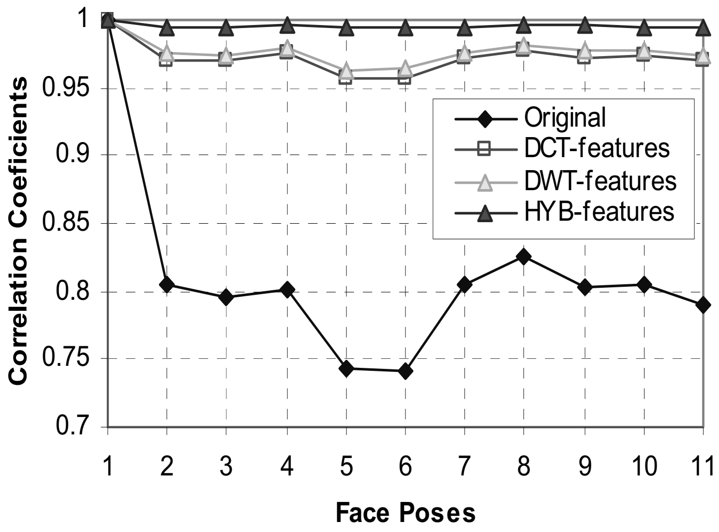


FIGURE 3.2: The correlation coefficients comparison between the original image and the compact face for eleven face poses of India face database

Its equation is shown below:

$$\begin{aligned} \|Q,T\|_q &= w_{0,0}|Q[0,0]-T[0,0]| \\ &\quad - \sum_{i,j:\overline{Q}[i,j]\neq 0} w_{i,j}\{\overline{Q}[i,j]=\overline{T}[i,j]\} \end{aligned} \tag{3.3}$$

Where $\overline{Q}[i,j]$ and $\overline{T}[i,j]$ are the query and target face features. The term $\left(\overline{Q}[i,j]=\overline{T}[i,j]\right)$ will return 1 if $\overline{Q}[i,j]=\overline{T}[i,j]$, and 0 if otherwise. Because hybrid face features consist of statistical and frequency features, a modified multi-resolution metric equation will be used to compute the similarity score, as written below.

$$\|F^Q,F^T\|_q = \sqrt{\frac{1}{4}\sum_{i=1}^{4}\left(v_i^Q - v_i^T\right)^2} - \sum_{j:\overline{f}_j\neq 0}\left(\overline{f}_j^Q = \overline{f}_j^T\right) \tag{3.4}$$

Where j=1,2,3, ... , m, v is statistical face query (i.e. μ and σ sub-vector), f is frequency features (i.e. f_i^{DCT} and f_i^{DWT}), Q and T represent to the query and target features. The smallest score is conclude as the best likeness (matching criterion). The equation 3.4 only needs subtraction and comparison operations. Therefore, the matching process that is defined by equation 3.4 is fast process because most

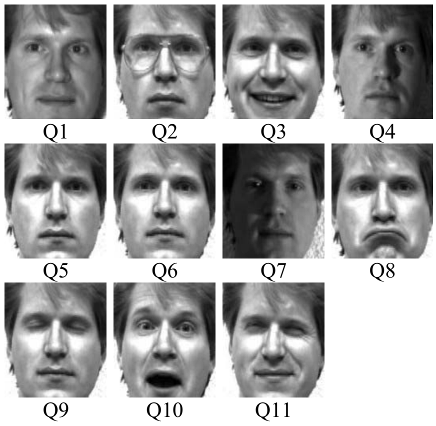


FIGURE 3.3: Example of the YALE database pose variations.

of the operations are comparison. Computer systems perform comparisons faster than subtractions or multiplications operation.

3.2 Experiment and Result

The experiment setup of this system was the same as that of the Chapter 2. However, we added one more well known YALE face database [49]. The Yale database is grayscale database, which consists of 15 people, and each person consists of 11 differential face expression, illumination, and small occlusion (by glass). Therefore, there are 165 face images in the database. An example of the face pose variation of the YALE database is shown on Fig. 3.3.

The first experiment investigated the effect of hybrid features size on success rate. This was performed in the ITS-Lab. database, 1 training face per person, and with the ratio between DCT and DWT coefficient of 50%:50%. The test results are shown on Fig. 3.4. The success rate slightly decreases when the size of hybrid features increase. The hybrid features size which is sufficient to achieve a good success rate ranges between 64 and 160 elements.

The second experiment was performed to evaluate the stability of hybrid method to face pose variations compared to the Eigenface method. In this case we implement the Eigenface method that was proposed by Ref. citepeigFR. This experiment was performed on the India face database, one training face per person, and features of 64 elements. The result shows that hybrid features give better and more stable success rates than single features (i.e., DCT or DWT feature only) and Eigenface features for the tested face pose variations, as shown in Fig. 3.5. It means that DCT and DWT, as the main element of hybrid features extraction, can accurately extract face texture information. These results can be achieved because hybrid features contain good frequency resolution of face texture obtained by DCT, and good time-frequency resolution of the face obtained by DWT. It this case, the DWT covers the weakness of DCT in signal representation.

The third experiment was carried out to investigate the robustness of the proposed method. It used the four databases, one training face per person, and features of 64 elements. The experimental result shows that the proposed method is robust for the tested face databases, which the difference of success rate among four databases is not more than 2%, as shown in Fig. 3.6. Also, note that the proposed method is superior to the Eigenface method, because the proposed method uses the frequency content as features of face image, while Eigenface uses statistical features, which are obtained by analysing the whole face pixel intensity. Frequency content does not depend on pixel intensity but depends on the spatial (pattern) of the object.

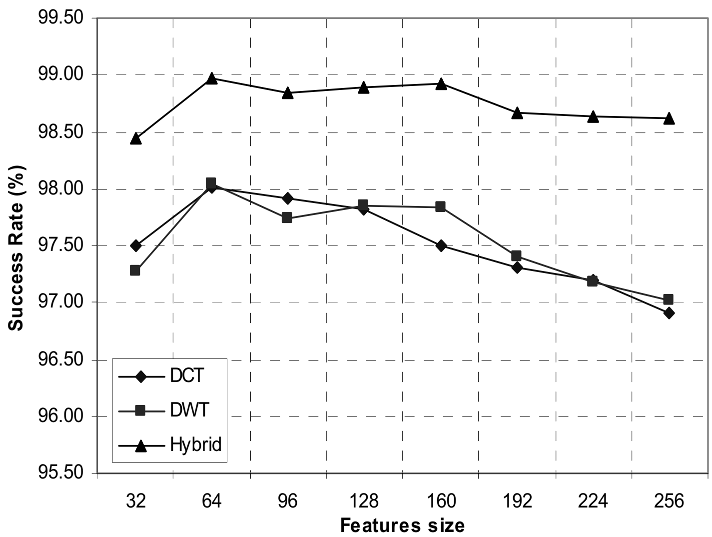


Figure 3.4: The effect of hybrid features size on success rate

The time consumption of the proposed method is shown in Fig. 3.7. It shows that the proposed system needs shorter time in both training and querying than Eigenface method. Our method reduced the training and query times of the Eigenface method by about 48.09% and 31.91% respectively while increasing the success rate by about 4.64%, when testing on 100 image gallery set. Tests were performed on

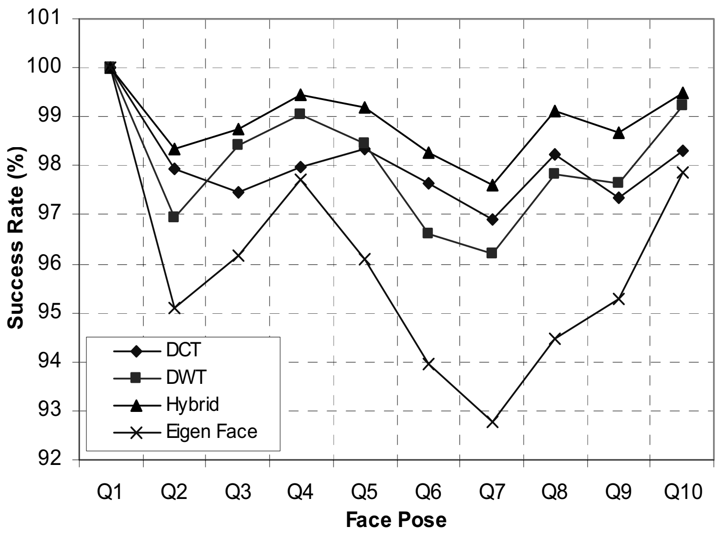


FIGURE 3.5: Robust comparison of Hybrid and other methods to face pose of the ITS-Lab. face database.

a personal computer with an Intel Celeron D 3.06 GHz, 1 GB RAM, and 160 GB hard disk.

The fourth experiment was performed to examine the effect of the number of face training per class to success rate. It was performed in the ITS-Lab. face database and features of 64 elements. The result shows that the more face training per class is given, the better success rate is gotten, as shown in Fig. 3.8. The most significant success rate improvement is given when one face training per class is given and then decreases gradually until the worst one. It can be achieved because the more training face is given, the optimum statistical features is gotten and the more frequency features is used to represent the face class. However, if more training is given, it will take more memory spaces because each face training is represented by m+4 coefficients. For example, two-face training per class will require almost twice memory spaces of one face training per class. However, they do not improve significantly the success rate (i.e. by about 0.59 and 0.49 of DCT and DWT success rate).

The fifth experiment was performed to investigate the effect of the geometrical normalization and localization of face image to the recognition rate. It was performed in the ITS-Lab. face database, 1 training face per person, and features size of 64 elements. In this case, the normalization and localization was prepared manually based on eyes position. The result shows that the recognition rate with

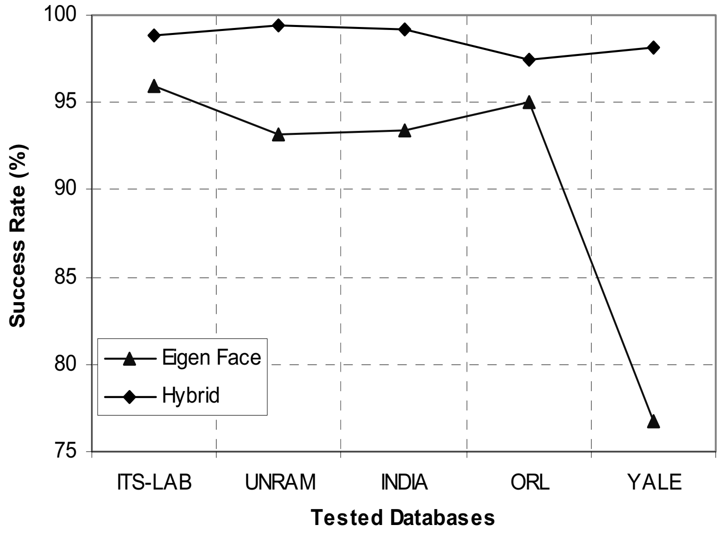


FIGURE 3.6: Success rate comparison of Hybrid and PCA methods on five face databases.

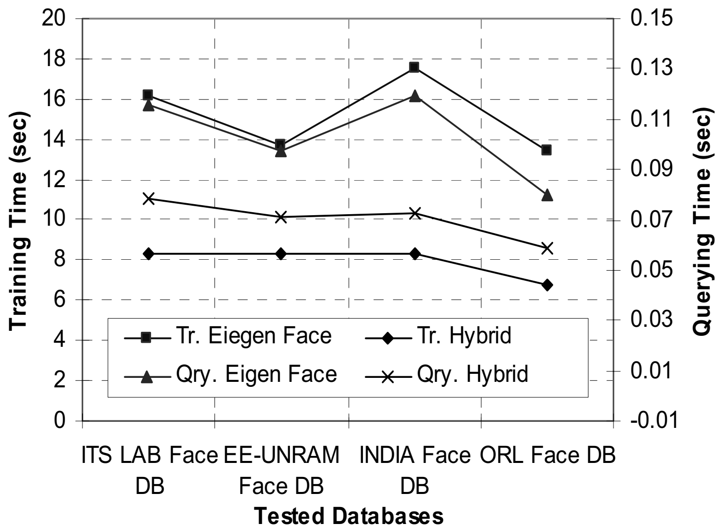


FIGURE 3.7: Training and query times comparison of two methods on four face databases.

normalized is better than that of without normalized. However, we have to create an automatic geometrical normalization and localization and these processes will increase the time consumption while not increasing the success rate significantly (just increase by about 0.15 point of non-normalized recognition rate), as shown in Fig. 3.9. In the next research, we will design the optimum geometrical normalization and localization for improving the proposed method's performance.

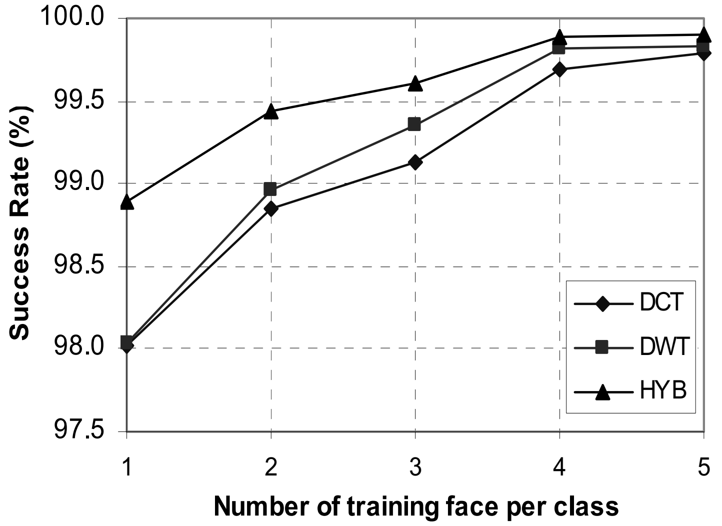


FIGURE 3.8: Success rate as function of number of face training per class.

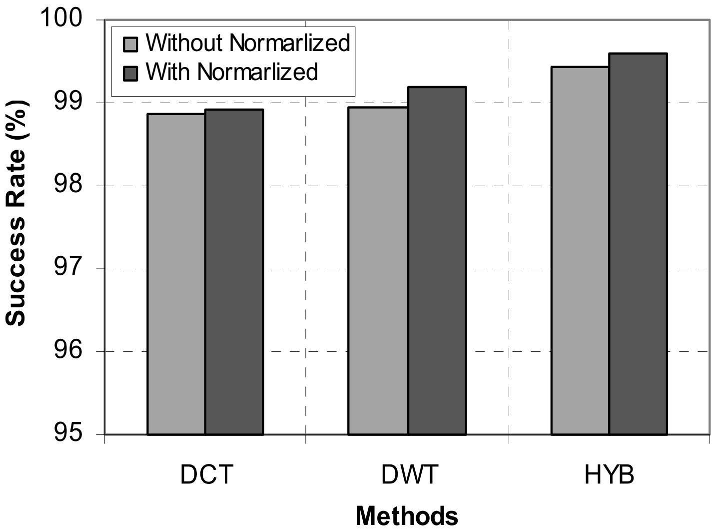


FIGURE 3.9: The performance comparison of normalized and non-normalized face image.

The last test was performed to know the accuracy of the proposed method compared to the others method (i.e. single face features (DCT and DWT) and PCA-based in the DCT and DWT faces features). In this test, we investigated two important parameters, namely False Rejection Rate (FRR) and False Acceptance Rate (FAR). The FAR is the success probability of unauthorized user to be falsely as accepted or recognized as legally user. The FRR is the success probability legally registered user to be falsely rejected by the system. If the value of the FAR and the EAR is equal then this point is called as equal error rate (EER).

The system that performs perfect classification or has high accuracy, is denoted by 100% true positive rate and 0% false positive rate or the value of EER is small or close to zero. We use ROC as performance evaluation because the ROC is not so needy on the precise variety of test data compared to confusion matrix and it is robust with respect to class skew [50].

The test was done on five databases (i.e., ITS-Lab, INDIA, UNRAM, ORL, and YALE). Those of databases are subjected as predicted positive (known face) and the frontal face image of CVL. database was subjected as predicted negative (unknown face). In this case, we added a threshold for the distance score to verifying input data whether is rejected or accepted. In other word, when an unknown face or a "claimed" identity is sent to the system for verifications and if the smallest distance of the face features of input face image among those of database images is less than the given threshold, the input face image is accepted and otherwise it is rejected.

The experimental results were plotted using receiver operating characteristics (ROC), as shown in Fig. 3.10. The results show that the proposed method performs better than other methods i.e. DCT, DWT, PCA+DCT, and PCA+DWT in all face databases which is denoted by the smallest ERR. In other word, the proposed method gives high enough accuracy, which is indicated by the average of EER about 0.0287. Moreover, because the experiments are performed on five faces databases, which have many variations of data as explained in four-first paragraphs of this section, and give high enough average success rate (about 97%, Fig. 3.6), small average EER (about 0.0287, see Fig. 3.10). It means that the proposed method has good enough reliability for the tested face databases. However, the ROC analysis in the YALE face database shows the worst curve for DCT and PCA method because of small number of face samples and one face image for training per class. The covariance matrix of training faces is almost singular when the number of face samples is quite small and it makes the projection matrix of the PCA is not optimum.

The proposed method is not compared with LDA methods because LDA-based face recognition must be trained with more than one face image per class [24, 28], while most of the experiments have been performed using one training face image per class. In addition, we do not concern about the face expression; however, the YALE face database consisting of large face expression is included in the experiment because the face features of any face pose is almost identical with the training face features without any preprocessing. The experimental result shows good enough performance because the YALE is small size face database. In the future work, we will consider the face expression in large face database using frequency analysis, which implement further face expression analysis such as face expression clustering using LDA method.

3.3 Discussion

The proposed method has been successfully implemented and shows good performance, robustness, stability, and accurately without requiring geometrical normalization. It can be achieved because hybrid face features have substantially good resolution of low-frequency components and good resolution in the time-frequency representation of face images. In this case, the low-frequency components are sufficient for face image representation because most of the information in a signal can be found at low frequency components. If an image is transformed to the frequency domain and then the high frequency components are removed, the reconstructed image will lose a little significant information. This phenomenon is successfully used for signal compression. In DCT analysis, the greatest information component exists in low coefficients indices (i.e., top left region) while in the DWT domain the low-frequency component exists in the approximation coefficients. Moreover, the DCT and DWT properties have advantages in features extraction, such as the ability to separate information signal into low frequency components uncorrelated with their counterpart frequency index. The DCT coefficients only provide frequency component information without localizing the specific frequencies in space. The DWT solves this problem by analysing the signal at different frequencies and at different times.

Hybrid dominant frequency features is an efficient way for reducing the original data dimension. The proposed method can reduce features by about 99.61% of original data size (i.e., 64 elements of 16384 elements). Therefore our method needs a short training and query time. Moreover, the hybrid features vector built using the fast-DCT (based on FFT) and the DWT, as shown in sub section 3.1. Computational complexity of the fast-DCT is N log2 N operations and the computational complexity of DWT is linear with the number (N) of computed coefficients (O(N)), where N is data dimensional. As note that the most common operation performed in our matching process is comparison, while Euclidean distance, which is implemented in the Eigenface method, performs mostly multiplication and addition operations. Furthermore, the computational complexity of the Eigenface method is $O(d^3)$ where d is the number of pixels in the training image as reported in Ref. [11].

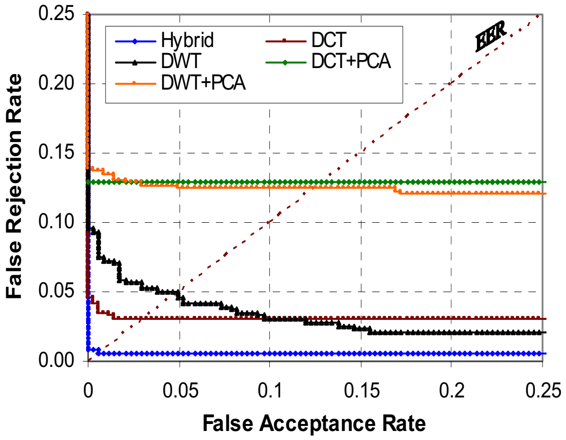
3.4 Conclusion

Multi-pose face image recognition based on hybrid dominant frequency features has been successfully implemented, which give good performance, robust to face pose variations, high enough a success rate (greater than 97%), and small enough EER. In addition, our purposed method has low computational cost, requires little memory space, and can overcome retraining problem. Based on these results, we think that our method can be implemented for real-time face recognition. However, this system still needs some improvements, such as applying cluster face analysis to make group of face based on skin color detection and implement the moment to detect the angles of faces capture.

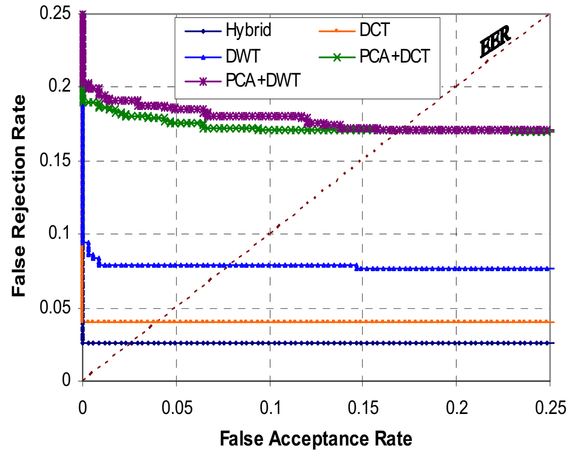
In the future, the research will focus on design of optimum geometrical normalization and localization, clustering analysis, and moment analysis for creating the face features in order to improve the accuracy of face recognition. Moreover, we will test the improvement methods using large face database such as FERET database.

Acknowledgement

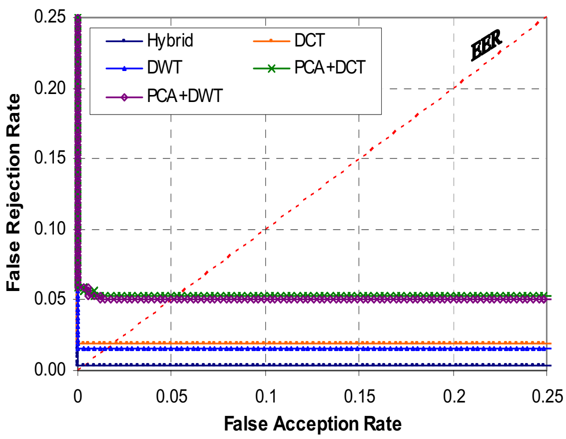
The authors gratefully acknowledge to Kumamoto university which supported our publication of this paper on IEICE Transactions on Information and Systems, to the owner the YALE face database. In addition, the gratefully acknowledge to IEICE reviewers who gave some helpful comments and suggestions for improving the presentation of this paper.



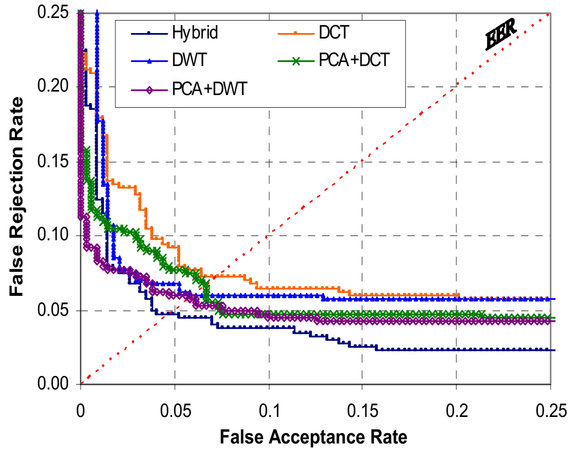
(a) ITS. Lab. database



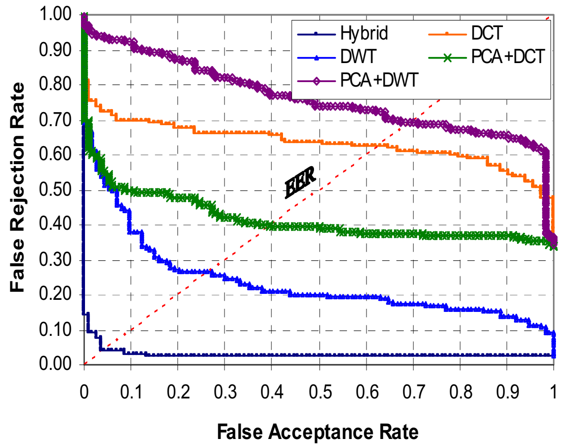
(b) INDIA database



(c) UNRAM database



(d) ORL. database



(c) YALE database

FIGURE 3.10: ROC curve of proposed method compared to the other systems.

Chapter 4

Statistical Based Face Recognition

4.1 Introduction

In the Chapter 2 and 3, we proposed an alternative approach to face recognition using quantized dominant frequency features which is called as compact features and hybrid compact features respectively. The performances in term of success rate is still far from the maximum (around 97%), because we did not apply features cluster for the classification. It means that the recognition was performed by only using dominant frequency discriminant information. However, our previous method has better performances than those of Eigenface (PCA) for both success rate and time processing. Further more the proposed method does not require the retraining process because features extraction does not depend on the global mean, as done by the PCA and LDA.

In order to improve the performance of face recognition, we propose an features cluster which is derived from maximum a posteriori (MAP) discriminant called as modified LDA (M-LDA). However, the propose features cluster works on non quantized dominant frequency features because that the quantized dominant frequency features make the global covariance be singular. In this case, the M-LDA is introduced to performed multi-pose face features cluster. There are several objectives of the our proposed method. Firstly, to redefine our previous compact and meaningful face feature without removing significant face image information.

Secondly, to build a simple classification technique that can classify face images to person's classes well. Thirdly, to make the M-LDA-based training system that can solve the retraining problem. Finally, to know the effectiveness of DCT and wavelet analysis for face feature extraction when they are combined with the M-LDA based classification.

The most related approach to our proposed system is face recognition based on LDA and its variations, as described in Refs. [8, 27, 28]. Ref. [27] proposed a combination of DLDA and FLDA called as RLDA to cover the weakness of classical LDA, in term of poor discriminatory and singularity problem. Ref. [28] implemented DCT to reduce data dimensional and only a small number of DCT coefficients were analysed by LDA. Ref. [8] implemented the wavelet transforms to reduce the dimensional of face image, employed a regulation scheme for the within-scatter matrix, and used optimization procedure. It was reported that the Daubechies (Db6) was implemented to filter image to resolution 29 x 23. However, those approaches still need high memory space and must be retrained when new face class is registered into the system. Ref. [24] proposed D-LDA to overcome the singularity problem LDA. However, Ref. [51] countered the claim that was presented in Ref. [24].

In our proposed method, we implement the frequency analysis (i.e. non-blocked DCT or wavelet transforms) for reducing the original data dimensional and the M-LDA for classifying the face class without localization and bounding box processing. It is difficult to compare time consumption of our proposed method with the existed works because the time consumption of the existed methods was rarely reported and the tests were carried out in different databases. Therefore, the results of our approach will be compared to the classical PCA, LDA, DLDA, RLDA, and SLDA, which will be tested with data from several challenges face databases.

This Chapter is organized as follows: section 4.2 describes brief PCA and LDA classifier and how they are implemented for face recognition; section 4.3 explains M-LDA based face recognition and its advantages and disadvantages, section 4.4 explains how to implement the frequency analysis as a main element of face feature extraction, section 4.5 describes the recognition algorithm of the proposed method, section 4.6 presents the experimental results and their discussion, and the rest concludes the Chapter.

4.2 Brief PCA and LDA Classifier

The main aim of the PCA and LDA is to find a transformation data such that feature clusters are most separable after the transformation. The most popular PCA analysis that is implemented for face recognition is Eigenface algorithm [18], as described below.

Suppose we have (X_1^1, x_1^1, C_1), ... , (X_{N1}^1, x_{N1}^1, C_1); ... ; (X_1^L, x_1^L, C_L), ... , (X_{NL}^L, x_{NL}^L, C_L), are images sample form L classes, each class has N samples, X_i^k represents image matrix of i-th samples of k-th class, x_i^k represents features vector of i-th samples of k-th class, where i=1, ... , N_k, and N_k is number of training samples of class C_k. Let $N = \sum_{j=1}^{L} N_j$ is total samples size. Let define μ_k as mean features vector of class C_k, and μ_a as mean features vector of all samples: $\mu_k = (1/N_k)\sum_{i=1}^{N_k} x_i^k$ and $\mu_a = \sum_{k=1}^{L}(N_k/N)\mu_k$. From the data sample, we can calculate the global covariance of features vector

$$C_g = \frac{1}{N}\sum_{k=1}^{L}\sum_{i=1}^{N_k}(x_i^k - \mu_a)(x_i^k - \mu_a)^T \tag{4.1}$$

The classical PCA (CPCA) projection matrix W can be obtained by eigen analysis of the covariance matrix C_g using the following equation:

$$C_g w_i = \lambda_i w_i, i = 1, 2, 3, ..., m \tag{4.2}$$

Where w_i is the i-th eigen-vector and λ_i is the i-th eigen-value of C_g. The projection matrix will be optimum if the projection matrix, W, satisfies the following criterion[2]:

$$J(W) = arg\max_{W}(W^T C_g W) = trace(C_g^P) \tag{4.3}$$

This criterion can be fulfilled by selecting m orthonormal eigen-vectors of C_g corresponding to the largest eigen-values (i.e. $m < n$), then placed them into $W = [w_1, w_2, w_3, ..., w_m]$ called as optimum projection matrix. Next, the projected features vectors can be obtained by equations below:

$$y_i^k = W^T x_i^k \tag{4.4}$$

The same as the PCA, the LDA determines the optimum projection matrix from both between-class scatter matrix S_b and the within-class scatter matrix S_w , as explained in Ref. [23, 28]. The optimum projection matrix (W) of LDA has to satify the following criterion.

$$J(W) = arg\,\max_{W} \frac{| W^T S_b W |}{| W^T S_w W |} \tag{4.5}$$

Where S_b and S_b are defined as

$$S_w = \frac{1}{N} \sum_{k=1}^{L} \sum_{i=1}^{N_k} (x_i^k - \mu_k)(x_i^k - \mu_k)^T \tag{4.6}$$

$$S_b = \frac{1}{L} \sum_{k=1}^{L} P(x^k)(\mu_k - \mu_a)(\mu_k - \mu_a)^T \tag{4.7}$$

The $W = [w_1, w_2, w_3, ..., w_m]$ which satisfy the Eq.(4.5) can be obtained by solving eigen problem of matrix ($S_b^{-1} S_w$) and then select m orthonormal eigenvectors corresponding to the largest eigenvalues (i.e. $m \ll n$). Finally, the optimum projected data can be determined by Eq.(4.4).

The main problem of the PCA method is lack of power discriminant while the LDA has the singularity problem of scatter matrix due to the high data dimensional and small number of training samples called as small size problem (SSS). Furthermore, they require retraining of all samples to obtain the most favourable projection matrix because the C_g and S_b depend on the global means. Regarding to the SSS problem of LDA, some methods have been proposed to solve that problem, such as DLDA, RLDA, and PCA+LDA. However, those methods still require large computational costs, memory space requirement, and retraining problems.

4.3 Modified LDA Classifier

in order to avoid the retraining problem and performing eigen analysis of LDA, we proposes an alternative LDA called as modified LDA (M-LDA). The M-LDA works under an assumption that the matrix scatter has small dimension and the global covariance of the training images is multivariate normal distribution. The

LDA works using both within class scatter (S_w) and between class scatter (S_b) that they are multivariate normal distribution. By using only the S_w, we can classify of query face features (x_q) to person's class using the equation below.

$$F_c = max\left[g_1(x_q), g_2(x_q), g_3(x_q), ..., g_m(x_q)\right] \tag{4.8}$$

with $g_i(x_q)$ is given by:

$$g_i(x_q) = \mu_i^T S_w^{-1} x_q - 0.5\mu_i^T S_w^{-1}\mu_i \tag{4.9}$$

where μ_i is vector column representing the mean of i-th class and (S_w) is with in class matrix determined using Eq. (4.6). The Eq. (4.9) is derived from maximum a posteriori (MAP) discriminant, as described below:

$$\begin{aligned} g_i(x) &= P(\omega_i|x) = \frac{P(x|\omega_i)P(\omega_i)}{P(x)} \\ &= \frac{1}{(2\pi^{\frac{1}{2}})|C_i|^{\frac{1}{2}}} exp\left\{-\frac{1}{2}(x-\mu_i)^T C_i^{-1}(x-\mu_i)\right\} P(\omega_i)\frac{1}{P(x)} \end{aligned} \tag{4.10}$$

where $P(x)$ is total probability of x. By eliminating the constant term and taking the natural log, it can be simplified as.

$$g_i(x) = -\frac{1}{2}(x-\mu_i)^T C_i^{-1}(x-\mu_i) - \frac{1}{2}\log(|C_i|) + \log(P(\omega_i)) \tag{4.11}$$

As mentioned previously that all class has identical covariance (C) and all class has the same prior probability ($P(\omega_i)$). Consequently, the part of $\log(|C_i|)$ and $\log(P(\omega_i))$ become constant, therefore the equation (4.11) can be simplified as below:

$$\begin{aligned} g_i(x) &= -\frac{1}{2}(x-\mu_i)^T C^{-1}(x-\mu_i) \\ &= -\frac{1}{2}x^T C^{-1}x + \mu_i^T C^{-1}x - 0.5\mu_i^T C^{-1}\mu_i \end{aligned} \tag{4.12}$$

By keeping just the terms which depend on μ_i and C, then substituting the C with the within class scatter, S_w, and finally the equation (4.9) is obtained.

This algorithm has some advantages for classifying the face image class:

1. It is simple because it does not require the eigenvalues and eigenvectors.
2. It can solve the retraining problem. It can be illustrated: firstly, when a new class added into the system, the M-LDA just calculates the mean and the covariance of its class; secondly, the newest mean is placed into the matrix M; and finally, the previous covariance is updated by adding it with the newest class covariance.
3. The computation complexity is less than the that of PCA and LDA because of not requiring eigen analysis.

The weakness of M-LDA is singularity problem due to the high data dimensional when the raw image as the input of this algorithm and small number of training samples. To overcome the singularity problem, we implement frequency analysis as dimensional reduction, as explained in the chapter 2 and chapter 3. However, the m selected coefficients do not quantized into +1 and -1 but they are normalized by its norms.

4.4 Face Recognition Algorithms

The proposed face recognition algorithm can be illustrated briefly as Fig. 4.1. There are three main process in this algorithm, namely preprocessing, training, and recognition. The preprocessing unit consists of color space transformation, equalization, and frequency analysis. In this research, NTSC (YIQ) color space is implemented for to convert the color image (RGB) to gray component (Y). Actually, there are two main function of preprocessing process: firstly, to decrease the effect of the non-uniform lighting condition on image face, which is performed using standard equalization, secondly, to create compact face features, using DCT or multiresolution wavelet transforms as described in the next sub-section. In the training process, the compact face features set is analysed by M-LDA for finding mean of each class and the within class covariance and then save them in database as a meaningful data for face classification.

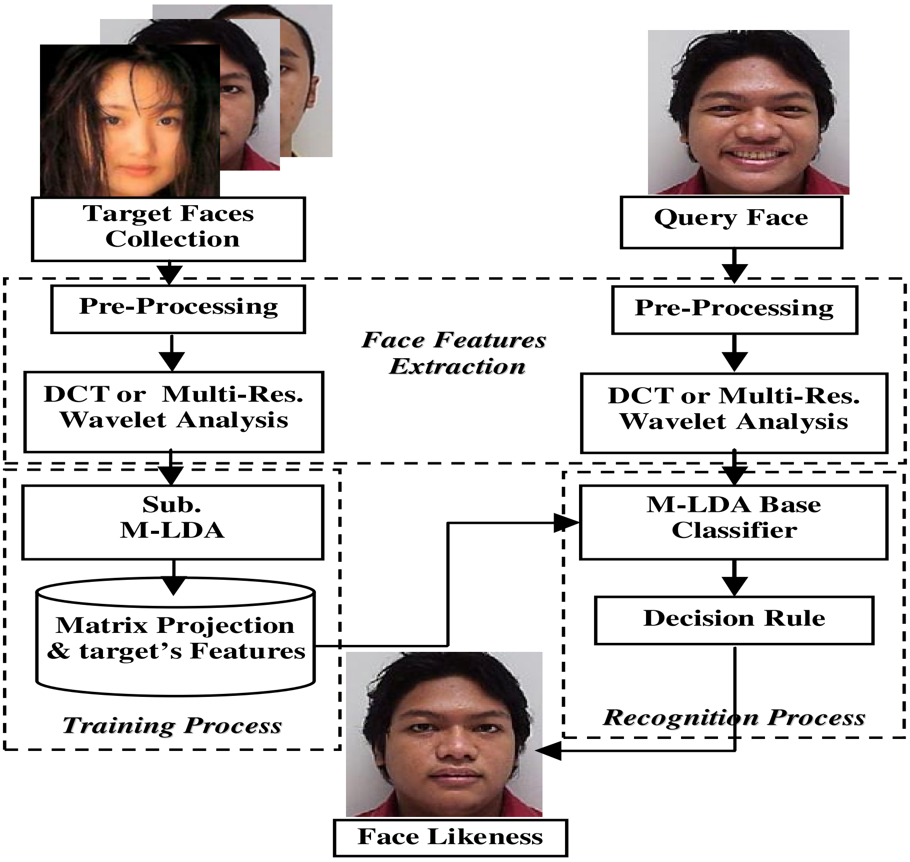


FIGURE 4.1: The proposed face recognition algorithm.

Finally, we calculate the similarity score between the input face features and the training face features set using MLDA based classifier (Eq. (4.8) and Eq. (4.9)). In this case, the maximum score is concluded as the best likeness.

4.4.1 Multiresolution Wavelet Analysis

In this research, we will investigate the DCT and multiresolution wavelet analysis based global features as dimensional reduction of raw image. This approach is different from our previous approaches presented in Chapter 2 and 3 in term of quantization and wavelet analysis. Previously, we implemented quantized dominant frequency content as holistic face features while in this chapter, we use non-quantized dominant frequency content of DCT and multiresolution wavelet analysis as holistic face features. For the DCT decomposition is the same as the previous one but for the wavelet is different in term of getting more detail information.

The multiresolution wavelet analysis is performed by implementing repeatedly classical DWT called as filter bank decomposition, as shown in Fig. 4.2. The A, H, V, and D are calculated by equation (3.4).

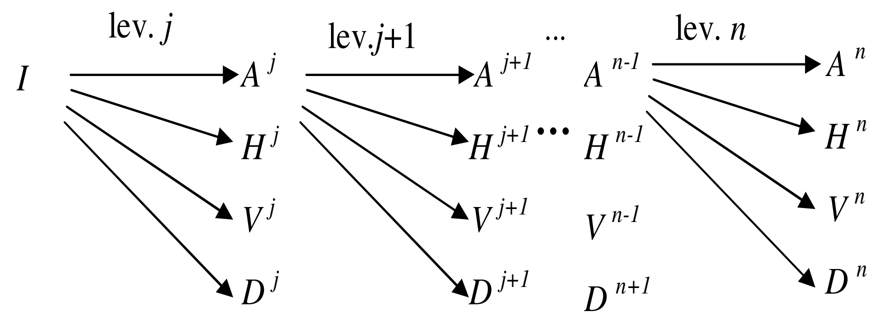


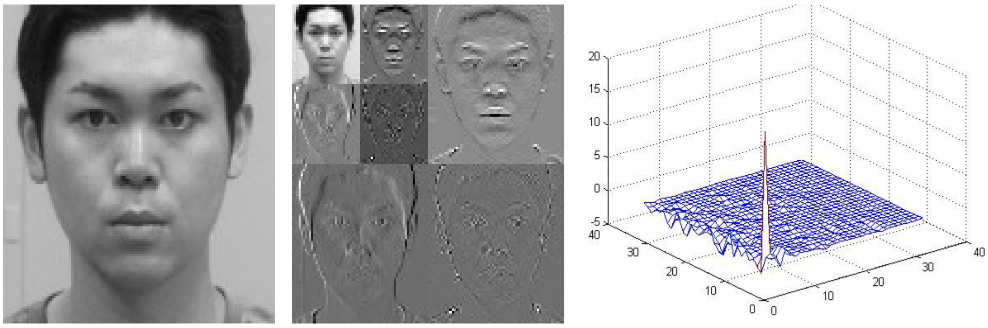
FIGURE 4.2: The proposed face recognition algorithm.

In order to make simple and fast decomposition process, we apply two different Daubechies wavelets basis, namely Db4 and Db1. First, Db4 basis decomposes face images until level 2 and it just returns the approximation coefficients. Second, the Db1 basis decomposes the Db4's approximation coefficient until maximum level. The pseudo code of multiresolution wavelet analysis is written below.

```
func dwtMultiDecompose(I:array[0 r-1, 0 .. c-1])
  // First step for DB 4 basis
  for res = 1 to 2 do
    [A, H, V, D]_res = dwt(I, h, g);
    I = A_res;
  end for
  // Second step for DB1 basis
  j = size(A,1)
  c ← A/2^j
  g ← 2^j
  while g ≤ 2 do
    for row ← 1 to g do
      decStep(c[row, 1..g])
    end for
    for col ← 1 to g do
      decStep(c[1..g, col])
    end for
    g← g/2
  end while
    return (c)
```

end func

The multiresolution of wavelet coefficients are denoted by c. The pseudo code above will return the wavelet decomposition coefficient as shown in Fig. 4.3.

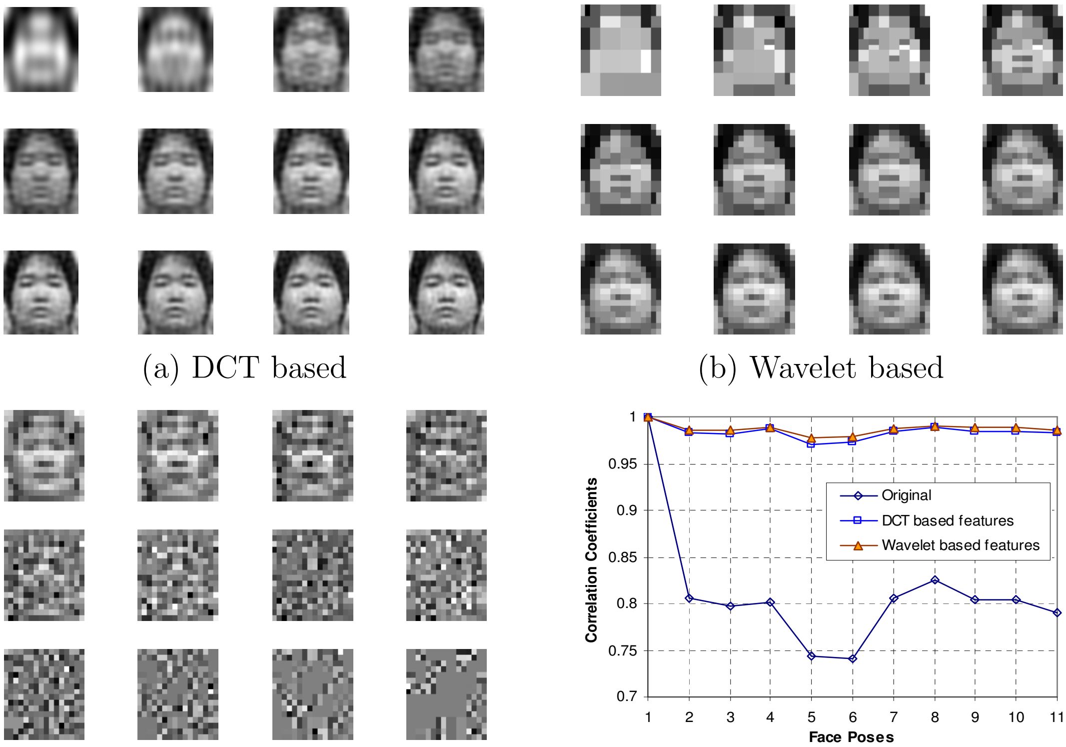


(a) (b) (c)

FIGURE 4.3: The output of multiresolution wavelet analysis pseudo code: (a) original image, (b) the first step decomposition coefficients, (c) the second step decomposition coefficients.

Each face image is represented using a small number of the DCT or wavelets coefficients, which are called as global face features. The global face features is a compact and meaningful information created by three steps, as done in Chapter 2 and 3 without quantization: firstly, convert the frequency domain coefficients to vector using row ordering technique; secondly, sort the vector descending using quick sort algorithm, finally truncate a small number of vector elements (i.e., less then 100 elements).

Consider the selected coefficients, if they are reconstructed into the face images, the reconstructed face images will be different. However, we can still understand that they are the face images, as shown in Fig. 4.4(a and b). Meanwhile, if the dominant frequency components (the selected coefficients) are removed the reconstructed face images are exactly different and we do not know that they are face images at all, as shown in the Fig. 4.4(c). This illustration proves that the most information of image exists in low frequency components. Furthermore, if the difference between the reconstructed image and the original is determined by root mean square error (RMSE), we will get results as shown in Table 4.1. This illustration proves that the most information of image exist in low frequency components.



(a) DCT based (b) Wavelet based

(c) Removed features (d) Correlation coefficients

FIGURE 4.4: The reconstructed images of the compact features which the size ranges from 16 elements until 225 element of the 16384 elements (original size) and the correlation coefficients comparison between the original image and the compact features of eleven face poses of India face database.

TABLE 4.1: The comparison of the RMSE of reconstructed to original image.

Frequency Analysis	RMSE as function features size						
	16	25	36	49	64	81	100
DWT	6280.2	3964	3301.6	3300.9	**1874.6**	1468.2	1257.4
DCT	2796.2	2581.2	1120.1	1007.8	**847.3**	1317.7	1190.1

Next, the compact face feature can be used as a basis of multi-pose face recognition because the compact face feature consists of dominant of frequency components of the face image. In this case, frequency analysis (i.e DCT or DWT) transforms the face information of any face pose variations in a single face to the collocation close to identical holistic information. It can be proved by calculating the correlation coefficients of both compact face feature and original image of any face pose variations in a single face, as shown in Fig. 4.4(d). It shows that the correlations of compact face features are almost the same for all face poses, however, the correlations of the original have large different values for all face poses.

4.5 Experiments and Results

The experiments were performed in several challenge face databases: the ITS-Lab. Kumamoto University database, the EE-UNRAM database, the India database [53], the ORL database [46, 47]. We used two main parameters to justify our proposed method performance: cumulative recognition rate in term rank of 4-th (recognition rate) and equal error rate of receiver operating characteristics (ROC) curve. All of the experiments were done in personal computer with specification Intel Celeron D 3.06 GHz processor and 1 GB RAM.

4.5.1 Investigate the performance of M-LDA)

In order to know the performance of M-LDA some experiments were performed as described below. The first experiment investigated the effect of face vector size to

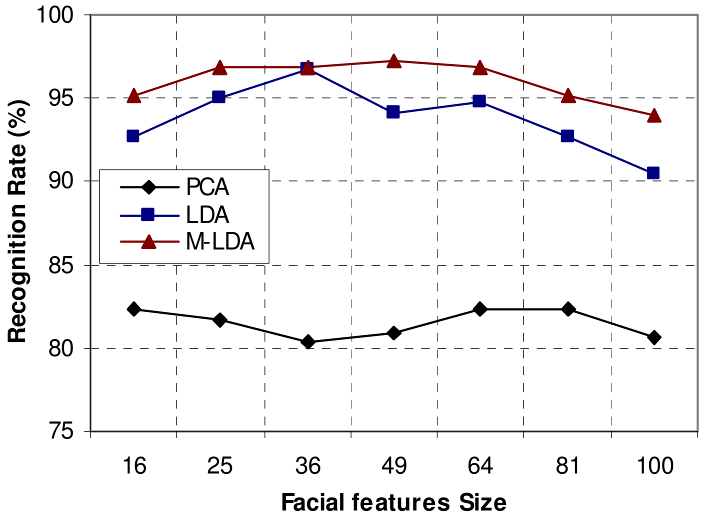


FIGURE 4.5: The effect of hybrid features size on recognition rate

recognition rate. This test was performed in the ITS-Lab. database, DCT-based feature extraction, and 4 training faces per class. The recognition rate slightly decreases when the size of face feature increases, as shown in the Fig. 4.5. The face feature size that is sufficient to achieve a good recognition rate ranges between 25 and 64 elements. This result can be achieved because the most face information exists in low frequency components (proved in Fig. 4.4(a,b)), therefore the face can be distinguished by just analysing them. In addition, face images of ITS-Lab were captured by good camera and the face pose variations were not very large, as shown on Fig. 7(a).

The second experiment was performed to evaluate the stability of the proposed method against face pose variations. It was performed in the India database, 4 training faces per class, DCT-based feature extraction, and face feature of 36 elements. The result shows that our method gives better and more stable recognition rate than classical LDA and PCA, as plotted in the Fig. 4.6. This result can be achieved because DCT has good energy compactness, therefore an image information can be represented by little mid-range frequency components. Furthermore, the image information of any face pose variations in a single face is almost identical (see Fig. 2.6(a)).

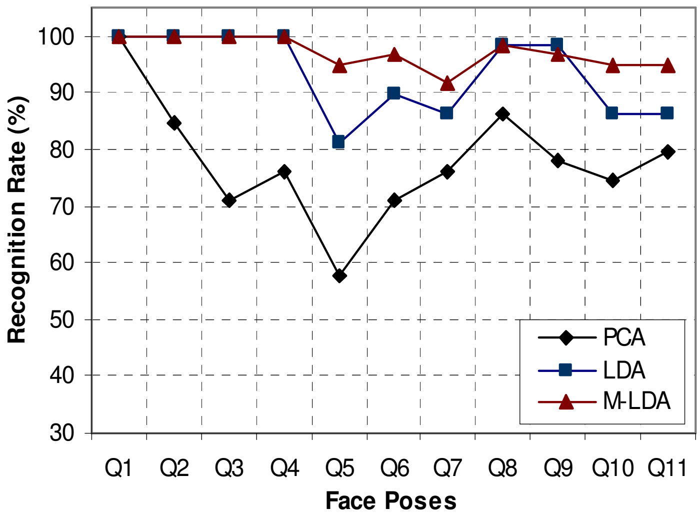
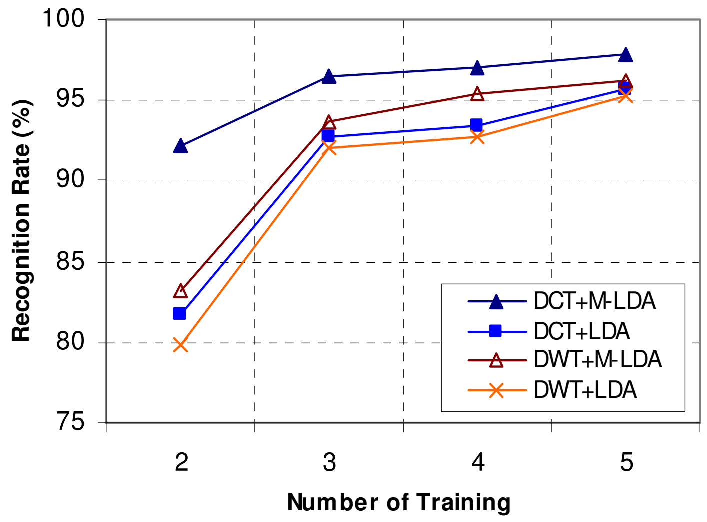


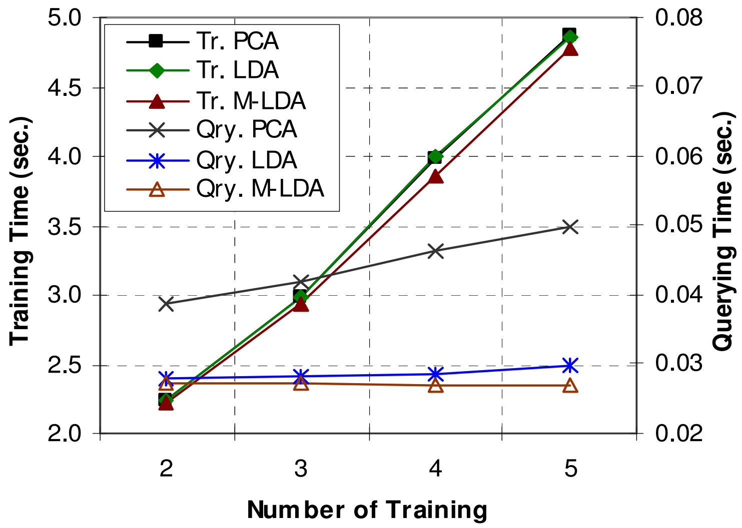
FIGURE 4.6: The effect of hybrid features size on recognition rate

The third experiment was carried out to investigate the effect of number of training face to recognition rate. It was performed in the India database, and face feature of 36 elements. We also compared the DCT-based with DWT-based face feature extraction. The result shows that the more training face of each class is given, the better recognition rate is achieved as plotted in the Fig. 4.7(a). It means that the more face training is given, the better mean and covariance matrix is gotten. The result also shows that the DCT analysis is more powerful than DWT analysis.

The time consumption of the proposed method is shown in the Fig. 4.7(b). It shows that the proposed system needs less time for both training and querying than those of PCA and LDA method. Our method reduced the training and



(a) Recognition rate



(b) Training and querying times

FIGURE 4.7: The effect of number of training face to the performance of the proposed methods.

querying times of the classical PCA method by about 17.37% and 37.87%, the classical LDA method by about 1.86% and 5.40% respectively while increasing the classical PCA and classical LDA recognition rate by about 17.37% and 1.78% respectively.

The fourth experiment was carried out to investigate the robustness of the proposed method. It was performed with the data from four face databases, 4 training faces per class, and feature of 36 elements. The difference of recognition rate of the purposed method among four face databases is not more than 1%, as shown in the Fig. 4.8. It means that the proposed method is robust for the tested face databases. Furthermore, the proposed method is superior to classical PCA and the DCT analysis gives better performance than the wavelet analysis.

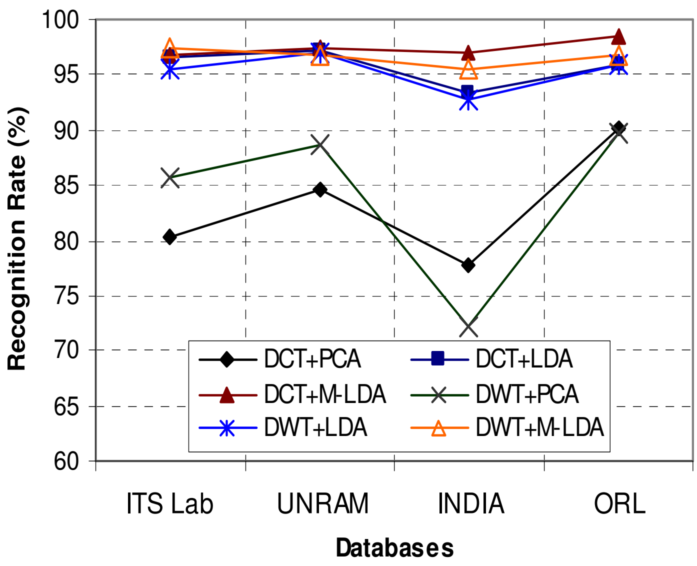
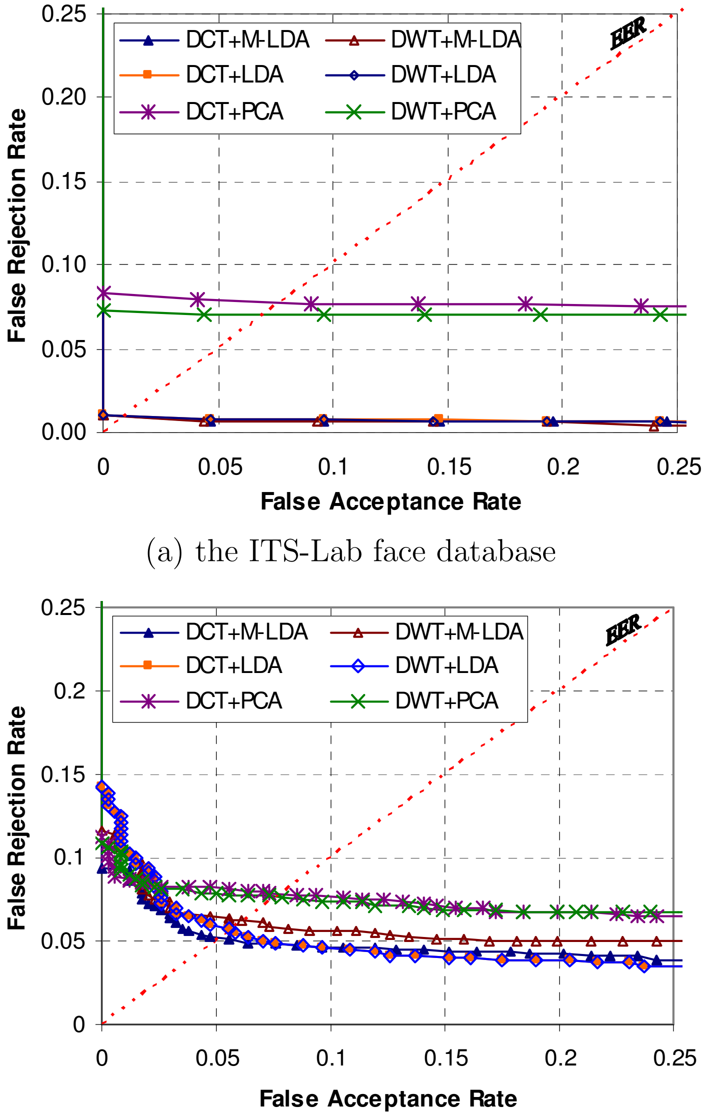


Figure 4.8: The comparison of recognition rate in term of rank 3 of tested face databases.

The fifth test was performed to know the accuracy of the proposed method compared to the classical LDA and PCA method. In this test, we investigated three important parameters, namely the false rejection rate (FRR), the false acceptance rate (FAR), and equal error rate (EER). The tests were performed in the ITS Lab. face database and the combination face database (i.e. ITS Lab. and India, and ORL). Those of databases were subjected as predicted positive (known face) and the frontal face image of other database was subjected as predicted negative (unknown face). As done previously, we added a threshold for the distance score to verifying input data whether is rejected or accepted. The system, which performs classification perfectly, is denoted by 100% true positive rate and 0% false positive rate or the value of EER is very small or close to zero.

The experimental results were plotted using receiver-operating characteristics (ROC), as shown in the Fig. 4.9. The results show that the proposed method has almost the same performance with LDA in DCT and DWT domain but it is superior to PCA-based method in DCT and DWT domain. Even though the EER of our method is much the same with the LDA, however, our method need less training and querying times, as shown in Fig. 4.12(b) and can solve the retraining problem



(a) the ITS-Lab face database

(b) the combination face database

FIGURE 4.9: ROC curve of proposed method compared to the other systems.

which is proven in the next section. As note, we implemented the PCA-based face recognition as proposed in Ref. [18] and the LDA-based face recognition, as explained in the section 4.2, which the optimum matrix projection of PCA and LDA was obtained by selecting some eigenvectors that correspond to largest eigen-values.

Table 4.2: Time consumption comparison for non-gradual training.

Time Processing (seconds)	Methods			
	SLDA	DLDA	RLDA	M-LDA
Training	16.7340	12.7970	12.7650	12.6400
Querying	0.0335	0.0319	0.0322	0.0262

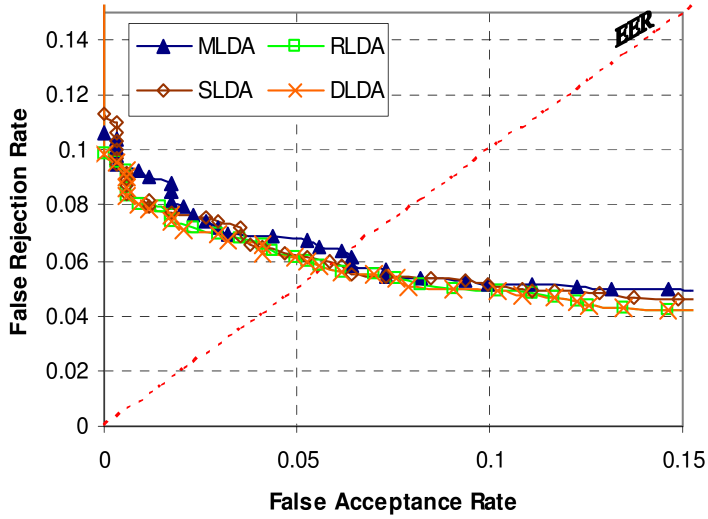
4.5.2 The comparison with established LDA

The next experiments were performed in order to know the strong point of the proposed method compared with the other established methods such as DLDA, RLDA, and subspace LDA. The tests were performed in all mentioned databases, 3 training face images per class, and face feature of 36 elements. The experimental results were plotted in the Fig. 4.9. The result shows that the proposed method gives much the same accuracy with DLDA and RLDA and little better than subspace LDA for both DCT and DWT domain face feature. However, the proposed method need less time processing for training and querying (Table 4.2) than those of them, because the proposed method does not require eigen analysis, as described clearly in section 4.3.

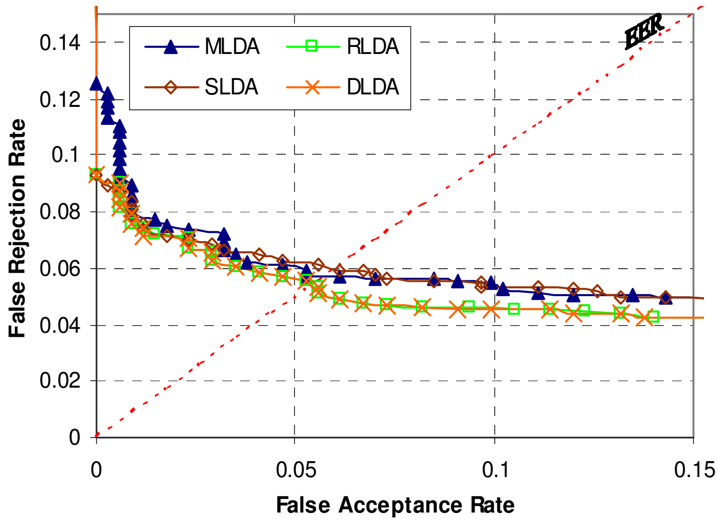
4.5.3 Retraining Solution

In order to show that the proposed method can overcome the retraining problem, the last experiment was performed. It was done in 160 face classes, 3 training face images per class, and face feature of 36 elements. The training was performed gradually, i.e. the first time, it was trained 40 face classes and then gradually added 10 new face classes to the system until 160 face classes. The experimental results were plotted in the Fig. 4.11-4.12.

The cumulative match score of the M-LDA method is strongly the same with other established LDA methods even though the M-LDA just considers within class scatter matrix for face class discrimination. It can be achieved because the M-LDA distinguishes the face class look like Mahalanobis distance. However, this method does not work at all for small sample size without implementing a compact face feature concept. Therefore, the face feature concept based on frequency analysis is the strength of our proposed method.



(a) DCT domain



(b) wavelet domain

FIGURE 4.10: ROC curve of proposed method compared to the other established method systems on the combination face database.

Regarding to the time consumption, the proposed method can solve the retraining problem of the LDA and PCA method. The proposed method needs very short training time when the training is performed gradually, as shown in the Fig. 4.12(a) and also requires less querying time, as shown in the Fig. 4.12(b). By these results, we prove the statement that M-LDA can crack the weakness of LDA and PCA in terms retraining problem. It can be done because the within class scatter matrix is obtained without considering the global means (mean of all samples) but just each class mean, as described clearly in section 4.3.

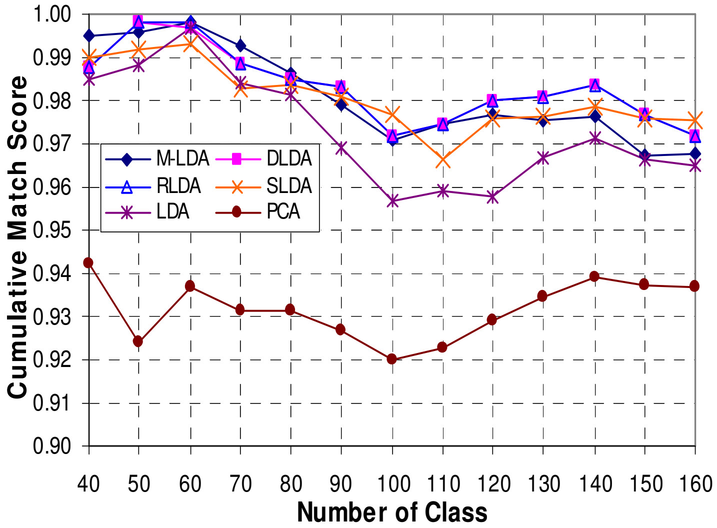


FIGURE 4.11: Cumulative Match Score (CMS) comparison for gradual training process.

4.5.4 Colour Information Consideration

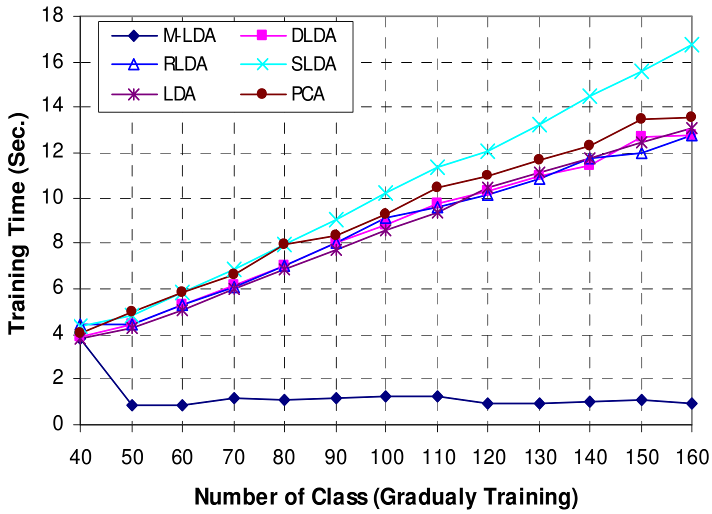
As known, the face images mostly cover with skin, which relate to the skin-tone color. Most of previous researches implemented YCbCr color configuration for modelling the skin tone color detection. From that model, it is known that the skin color can be found on the chrominance (*Cb* and *Cr*) components. In addition, the chrominance components of skin-tone color are non-linearly dependent on luminance [52]. Therefore, we include the *Cb* and *Cr* component of YCbCr color space transformation to cover the skin color information of face images.

In order to deal with the color information of face image, MLDA based classification has to be accessed three times which correspond to YCbCr color space then the class decision rule is performed by maximum-mean distance using the modification of equation (4.8).

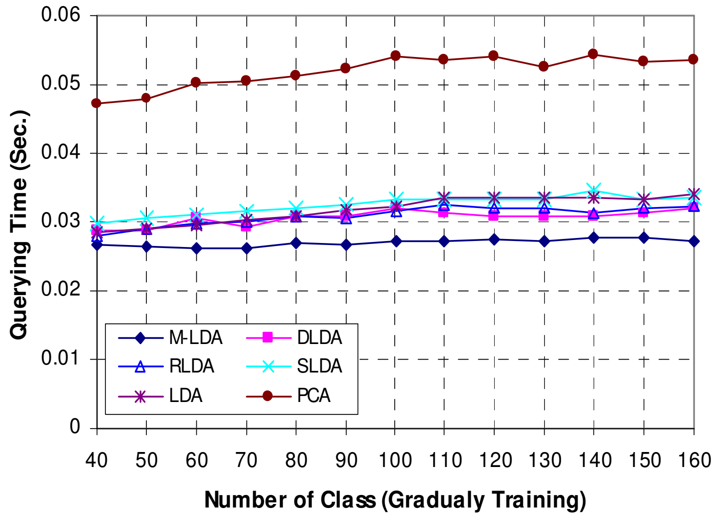
$$F_c^{Clr} = max\left[\bar{g}_1(x_q), \bar{g}_2(x_q), \bar{g}_3(x_q), ..., \bar{g}_m(x_q)\right] \tag{4.13}$$

where $\bar{g}_1(x_q)$ is mean of $(g_i(x_q^Y)$,$g_i(x_q^{Cr})$,$g_i(x_q^{Cb}))$.

To know the effect of the skin color information on the recognition rate, some experiments were carried out in the merger database which is combination of ITS-Lab. Kumamoto University database, EE-UNRAM database, India database, and Cropped face database [39, 40]. This database totally consists of 2268 images with



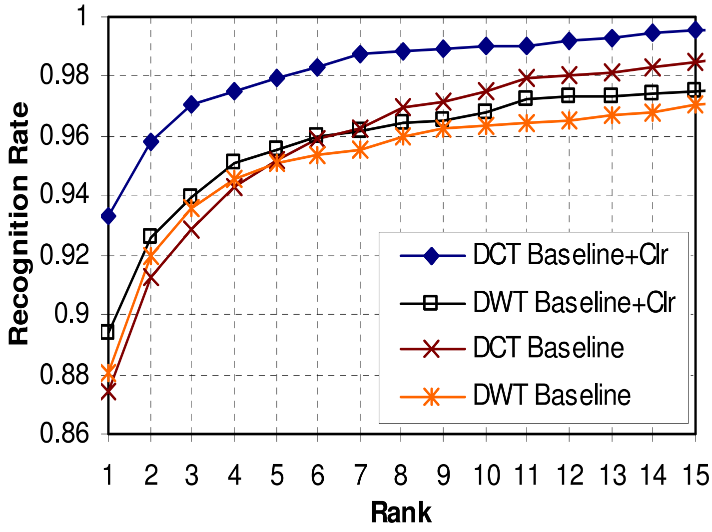
(a) Training process



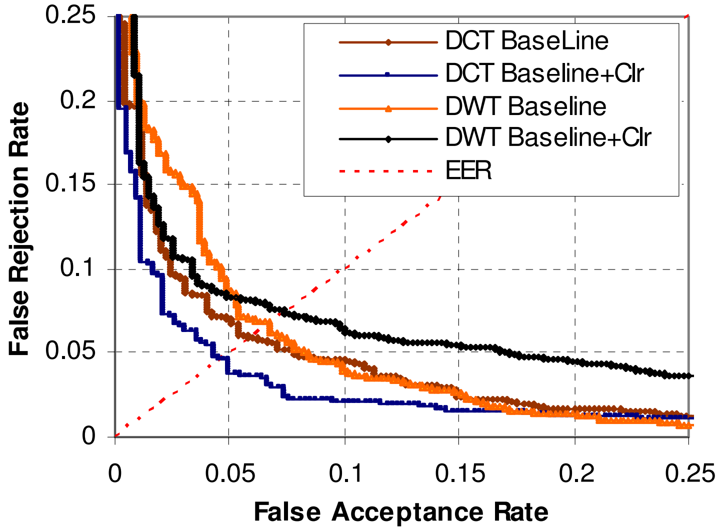
(b) querying process

FIGURE 4.12: Time consumption comparison gradual training process.

196 classes. From 2268 images, 1080 images are selected as training data and the remaining as query. The experimental result shows that the color information gives high enough improvement on recognition rate. In term of rank 1, the color information improves about 5.41% of the base line (MLDA in grayscale) and the color information improves the accuracy of base line which is indicated by lower EER (by about 0.0457 for DCT. baseline+Clr) than that of DCT baseline (by about 0.0585), as shown in Fig 4.13.



(a) CMS



(b) ROC

FIGURE 4.13: The face recognition performance of the grayscale and color configuration.

4.5.5 Discussion

The proposed method has been implemented successfully and showed good performance, robustness, and stability without requiring geometrical normalization. This performance can be achieved because some reasons:

1. The most signal information exists in low-frequency components, while the high-frequency components impart the flavour or nuance of signal. Consider the human voice, if we remove the high-frequency components, the

voice sounds different, but we can still tell what is being said. However, if we remove enough of the low-frequency components, we will hear gibberish. Therefore, the face image information can be represented by the low-frequency components.

2. Frequency analysis has good capability in separating the frequency components of the signal. By the DCT and DWT, the signal is decomposed to frequency matrix, which has good low-frequency resolution and good time-frequency localization.

3. The DCT and DWT coefficients uncorrelated with their counterpart frequency indices, it means that if most of non-dominant frequency component is removed it will not significantly decrease the quality of signal information.

4. The chrominance information tends to provide better result than that of previous method because the skin color information of face image is considered in the recognition system. It means that the skin color is important information of face image, which can be found on the chrominance components of color space transformation. Moreover, human being can visually distinguish face image class using just skin color

The compact face feature based on frequency analysis is a meaningful concept for multipose face recognition because the frequency analysis can extract the information of any face pose variations to a similar face feature. It can be done because the frequency of signal does not depend on its amplitude but its shape. Regarding to face image, even though the pose variations in a single face can be very large, however, visually they still have almost the same face shape. This phenomenon can be proved by correlation analysis, as shown in Fig. 4.4.

In addition, the frequency analysis is an efficient way for reducing the original data dimensional. By using proposed method, the face features can reduced by about 99.78% of the original size (i.e., 36 elements of 16384 elements), while the recognition rate is high enough with short time processing. The short time processing can be achieved by implementing the fast DCT (based on FFT) and DWT for face feature extraction. In addition, the M-LDA-based classification just requires the mean of each face class and the global covariance, while the LDA and PCA require global mean to obtain global covariance and eigen vectors for the projection. Theoretically, the computational complexity of the fast DCT decomposition is N log2 N complexity and the computational complexity of the DWT is linear with the number (N) of computed coefficients ($O(N)$), where N number of data. Therefore, our method needs short training and querying times.

4.6 Conclusion

A multipose human face recognition approach based on frequency analysis and M-LDA has been implemented successfully. Frequency analysis is an efficient way to create the compact face feature. Therefore, the compact face feature is good concept of reducing memory space requirement and computational load of LDA and PCA. M-LDA in frequency domain performs good enough classification of the persons class. Based on the experimental results, the proposed method provide better performances (i.e. higher recognition rate, more robustness, smaller EER, and shorter training and querying times) than those of LDA and PCA based face recognitions. Moreover, the proposed method can cover the retraining problem as explained in sub-section 4.3 and proven in section 4.5.3. Between the DCT and the DWT frequency analysis, the first one performs better for face image representation, which is shown by higher recognition rate and smaller EER than those of the DWT. Based on these results, we think that our method can be adopted for real-time face recognition. Furthermore, the color information can improve significantly the recognition rate of that of base line by about 5.41% in term of rank one. However, the color information requires longer time processing than that of the baseline.

This process needs some improvements, such as applying cluster face analysis to make group of face and implementing the moment to detect the angle of face capture. Next, the research will focus on clustering analysis and finding the optimum face classification method for frequency analysis in order to increase the previous performances.

Acknowledgement

The authors gratefully acknowledge to Kumamoto university which supported our publications in the Journal of the IIEEJ and ICEENG 2008 conference. Further more, the gratefully acknowledge to IIEEJ reviewers who gave some helpful comments and suggestions for improving my presentation of this chapter.

Chapter 5

Real Time Face Recognition Using Predictive LDA and Alternative PCA

5.1 Introduction

Face recognition is the most popular applications of image analysis and understanding and become one of the most active research areas. At least, there are two reasons for this trend: the first is the wide areas of commercial applications such as entertainment, smart cards, information technology, low enforcement, and surveillance; the second is the availability of feasible technologies of hardware and software nowadays. The face recognition is one to many matches which compare a query face features against all the training face features to determine the identity of a query face. The face recognition remains difficult to be done because variations in a single face can be very large, while the variations between different faces can be quite small. In addition to this, face variability also depends on ethnicity and registration technique (i.e., capture method, lighting condition, and devices). For instance: each ethnicity has special characteristic of face in term of skin and hair color, eyes and nose shape; the frontal face images show clearer information than that of lateral face images and so are the face images captured by high resolution camera than web camera. Moreover, other difficulty factors are high complex non-linear manifolds, high dimensionality, and small sample size.

The PCA, LDA and their combinations are most popular approach for features clustering, because of its uncomplicated processing. Typically, the PCA's optimum projection matrix is obtained by eigenvectors and eigenvalues analysis of global covariance matrix [1-5]. Meanwhile, The LDA's optimum projection matrix is obtained using eigen analysis of the within class (S_w) and between class (S_b) scatter[4-5]. Even though, the LDA and its variants provide better separable projected data than PCA but both of them still have some problems: large computational cost, large memory space requirement, retraining, and lack of separable projected data when the training data class size is large. In other word, the last problem is caused by overlap-projected features when the data class is large. The more data classes are, the more overlap-projected features will be.

In the previous chapter, the global/holistic features of face image which is based on dominant frequency content has been successfully implemented for face recognition. By using the global/holistic features concept as dimensional reduction of the face image can compress about 99.39% of the original size (i.e., less than 100 elements of 16384 elements) which gave good enough enough performance [33, 35, 38]. It means the frequency analysis based global features is an efficient way for reducing the original data dimensional. Therefore, that holistic features concept, which is combined with the alternative PCA (APCA), predictive LDA (PDLDA), and integration of them, is implemented to create real-time face recognition.

This proposed method is an alternative approach to face recognition algorithm that is based on global/holistic features of face image which is combined with APCA and PDLDA to overcome large computational costs and retraining problem of the conventional PCA and LDA and to improve features cluster using the APCA and PDLDA integration. There are several objectives of the proposed method. Firstly, to prove that the global features of face image contains most of face classification information. Secondly, to define alternative PCA and redefine the predictive between class scatter (S_b), which have the same structure as their original and less computation complexity. Thirdly, to optimize the recognition performance using multi-stage classifier integration. Finally, to know the effectiveness of proposed methods compared with the best traditional subspace methods. In addition, this proposed method will work on color face image instead of grayscale in order to cover the skin information. The skin color face image is one of discriminant information, which is available in the chrominance component and has to be considered to get good performance. Consequently, it will take longer time processing than just working in grayscale. Furthermore, shape face analysis performed by invariant moments is also included to get the holistic information of face pose variations.

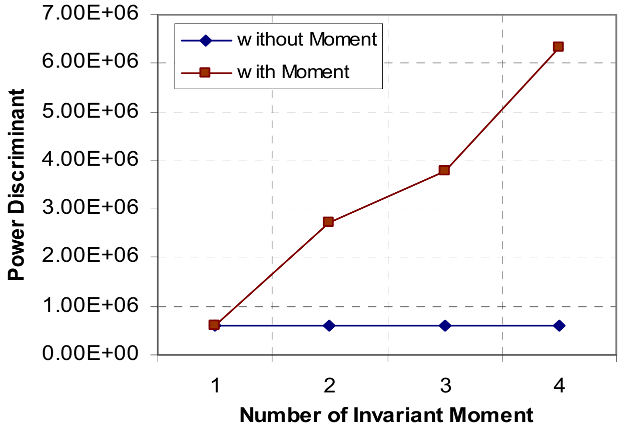
This chapter is organized as follows: section 5.2 describes color and moment feature extraction, section 5.3 explains briefly the algorithm of features cluster, which involves illustration of PCA and LDA, APCA algorithms , predictive LDA (PDLDA) algorithms, and multi stage classifier to improve the recognition rate achievement of our previous methods, section 5.4 presents the experiments setup and experimental results as well as the results discussion, and the rest concludes the paper.

5.2 Color and Moment Features Extraction

As we described previously, people can readily and accurately extract patterns. For example, people who have studied A character can perfectly recognize each style of the A character, even though the A character is written in other forms, scale, and color, such as **a**, **A**, ***A***, a, A, etc. In this case, the human being has an intrinsic technique to extract some pattern denoting an object's information.

As reported previously, the skin information that exist on chrominance component of YCbCr color space tend to provide good enough achievement has to be considered for holistic features to get as much as possible the discriminant information of face image. In order to get robust global face features of any face pose variations, the moment information that provides invariant measure of face images shape is considered. The moment information is obtained using invariant moment analysis, which is derived from central moment analysis [48]. The invariant moment set is invariant to translation, scale change, and rotation therefore this concept can be used to get the holistic information of any face pose variations. In this case, the invariant moment is determined just in the intensity component of color images. Finally, from both the moment invariant set (I) and the selected frequency coefficients (f), we construct a compact global features vector, $x_i = [f_i^{DCT}, I_i]$, where i is i-th class of face image. The dimension of x is m+n, where m is number of selected frequency coefficients and n is number of selected invariant moments.

As described and proved in the Chapter 1, 2, and 3, the dominant frequency content existing in low-frequency components is sufficient for face image representation. In other word, if an image is transformed to the frequency domain and then the high frequency components are removed, the reconstructed image will



(b)

FIGURE 5.1: The power discriminant as function number of invariant moment.

lose a little significant information (see Fig. 4.4). Furthermore, if the difference between the reconstructed image and the original is determined by RMSE, we will get results as shown in Table 4.1. It shows that the dominant frequency of DCT-based tend to contain most information of face image than that of the wavelet (DWT)-based face features.

The strength of the proposed global features: it has compact size, which has higher power discriminant than that of without moment information as shown in Fig. 5.1 and it has almost similar features of any pose variations of face images which has been proved in Ref. [16]. It means that the proposed global features can overcome the large variability of face pose variations in single face. In this case, the power discriminant was determined by Eq. (5.8). In addition, the invariant moments, which are higher than 4, are not included because they make the within class scatter matrix (S_w) be singular. This problem comes because those invariant moment's values are close to or event zero.

5.3 Features Clusters

5.3.1 Illustration of PCA and LDA

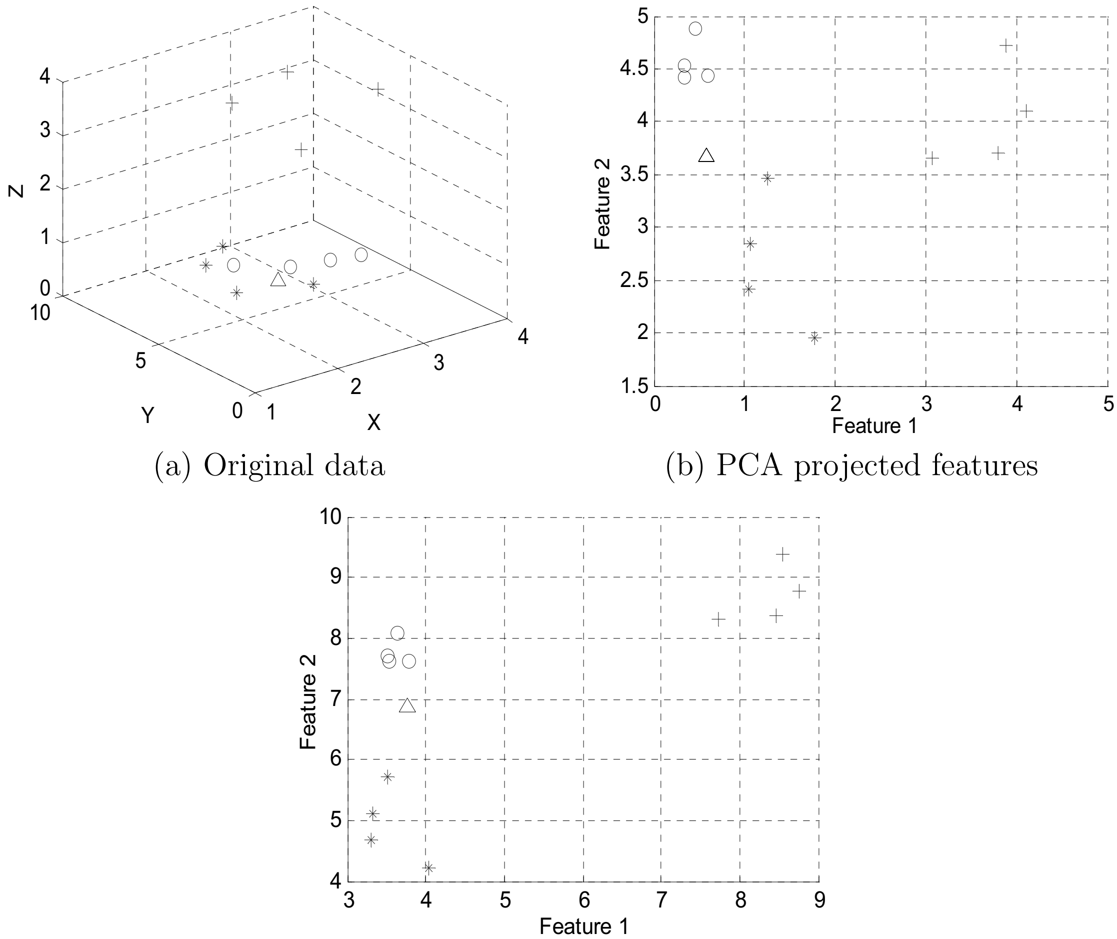
The aim of common PCA and LDA is to find a transformation data such that feature clusters are most separable after the transformation. The most popular

PCA analysis that is implemented for face recognition is eigen-face algorithm [1-3], as described below.

How the PCA and the LDA can performs the features clusters can be illustrated below. Suppose we have a training data set that is signed by plus (1-th), star (2-nd), circle (3-rd), and an initial data (triangle) which belongs to the circle class, as plotted in Fig. 5.2. It shows clearly that the star class overlaps with the circle one. In order to know how the triangle data overlap with other data classes and to know which class the triangle data belongs to, we compare two type methods: the PCA and LDA with weight score (WDLDA). Firstly, the training data set is projected using the PCA algorithms and then the projected data plotted as in Fig. 5.2(b). From Fig. 5.2(b), we still do not understand what the triangle data belongs to. If the euclidean distance of querying (initial) and training projected data are calculated, we will get: d = [0.7727 2.4954 0.7019]. Where *d(i)* is the distance between the querying class and i-th training data class, where i=1,2,3. From these distances, the querying data (Δ) is member of the third (star) class. This result does not match with the initial condition (the querying data is member first class (circle)). It means the projected data do not well separate yet.

In order to get well separable projected data, the weight-score is introduced. From LDA analysis, we get a S_w , S_b , and mean vector (μ_i) of each class. Because S_w is independent from the global mean of training data, the weight-score (wSc) can be determined using $wSc_i = \mu_i^T Sw^{-1} \mu_i$, i represents the data of i-th class. Each weight-score is introduced into each projected data then well-projected data is gotten as shown in Fig. 5.2(c). Then, if we calculate the distance of querying and training projected data it will give: d = [1.4113 5.6584 1.6941] scores. From these distances, the querying data (triangles) is member of the circle class (red). This conclusion is match with the initial condition. It means that the WLDA provides better separable projected features than that of PCA and LDA.

The main problem of the PCA method is lack of power discriminant while the LDA has the singularity problem of scatter matrix due to the high data dimensional and small number of training samples called as small size problem (SSS). Furthermore, they require retraining of all samples to obtain the most favourable projection matrix because the C_g and S_b are depends on the global means. Regarding to the SSS problem of LDA, some methods have been proposed to solve that problem such as DLDA, RLDA, sand PCA+LDA. However, those methods still require large computational costs, memory space requirement, and the retraining problems.



(a) Original data

(b) PCA projected features

(c) LDA projected features

FIGURE 5.2: Illustration of PCA and LDA projected features.

5.3.2 Alternative PCA

In order to get higher power discriminant and to solve retraining problem of the PCA, an alternative solution is proposed called as alternative PCA (APCA). The APCA does not work on global covariance (C_g, Eq. (4.1)) but on the within class scatter (S_w) of training set for getting the projection matrix. The optimum projection criterion of APCA is not the same as classical PCA (Eq. (4.3)) but using the following equation.

$$J_{APCA}(W) = \underset{W}{arg\,min} \mid W^T S_w W \mid \tag{5.1}$$

This criterion is derived form Fisher (the LDA) criterion Eq. (4.5). If the this criterion is called back without considering the between class scatter matrix (S_b)

or setting the S_b as identity matrix (I), then it can be written as

$$J_{APCA}(W) = \underset{W}{arg\,max} \frac{| W^T W |}{| W^T S_w W |} \tag{5.2}$$

Because the projection matrix is orthonormal ($W^T W = I$), therefore Eq. (5.2) can be simplified as

$$\begin{aligned} J_{APCA}(W) &= \underset{W}{arg\,max} \frac{| I |}{| W^T S_w W |} \\ &\cong \underset{W}{arg\,min} \, | W^T S_w W | \end{aligned} \tag{5.3}$$

In order to satisfy this optimum criterion, the determinant of the denominator should be as small as possible or even close to zero. It can be satisfied by selecting a small number of eigenvectors (m) of S_w corresponding to the smallest eigenvalue (i.e. $m < n$), then placed them into W.

This approach has two main advantages compared with the classical PCA (CPCA): it can solve retraining problem and its projected data is more separable, which provide higher discrimination power than that of CPCA. The first advantage can be achieved because the within class scatter matrix does not depend on the global mean of the training samples, which can be proved using the following illustration. Let recall the Eq. (4.6):

$$\begin{aligned} S_w &= \frac{1}{N} \sum_{k=1}^{L} \sum_{i=1}^{N_k} (x_i^k - \mu_k)(x_i^k - \mu_k)^T \\ &= \frac{1}{N} \sum_{k=1}^{L} S_w^k = \frac{1}{N} \left\{ \sum_{k=1}^{L-1} S_w^k + S_w^L \right\} \\ &= \frac{1}{N_{new}} \{ S_w^{old} + S_w^{new} \} \end{aligned} \tag{5.4}$$

where $S_w^k = \sum_{i=1}^{N_k} (x_i^k - \mu_k)(x_i^k - \mu_k)^T$ and $N = N_{new} = N_{old} + N_L$.

Let compare the Eq. (4.1) to Eq. (5.4), the Eq. (4.1) depends on global mean (μ_a), but the Eq. (5.4) does not do at all. If the new class data come to the system, the Eq. (4.1) has to recalculate the global covariance. It is contrast with the Eq. (5.4) which just calculates the newest class covariance, S_w^{new}, and adds it to the S_w^{old}.

The projected data of the APCA is most separable than CPCA because the *J(W)* of APCA is higher than CPCA. From Eq. (4.2), it is easy to understand that the eigenvalues of the global covariance or the within class scatter data is diagonal matrix, which the diagonal elements are the selected eigenvalues.

$$\lambda = \begin{bmatrix} \lambda_1 & 0 & 0 & \dots & 0 \\ 0 & \lambda_2 & 0 & \dots & 0 \\ 0 & 0 & \lambda_3 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \lambda_m \end{bmatrix} \tag{5.5}$$

substituting the Eq (5.5) into Eq. (5.1), then we get:

$$\begin{aligned} J_{APCA}(W) &= arg\,max \frac{\mid I \mid}{\mid \lambda^{APCA} \mid} \\ &= arg\,max(\mid \lambda^{APCA} \mid^{-1}) \\ &= arg\,max \left(\prod_{i=1}^{m} \frac{1}{\lambda_i^{APCA}} \right) \end{aligned} \tag{5.6}$$

Theoretically, the CPCA approach removes the null space (selecting m eigenvectors correspond to the largest eigenvalues), while the APCA approach keeps the null space of within class scatter (keeping m eigenvectors correspond to the smallest eigenvalues). It means one of the eigenvalues of global covariance or within class scatter is close to zero. Therefore, the determinant of APCA projected features covariance becomes very large or close to infinity compare to trace of CPCA projected features covariance. From this evaluation, we can conclude that the APCA's projected features vectors are more separable than those of CPCA.

Another optimum criterion can be used to justify the strength of APCA called as discrimination power (DP) that was proposed by Etemad, et.all (1997)[23], as written below.

$$S^W = S_w^{-1} S_b \tag{5.7}$$

$$J(W) = sep(W) = trace(S^W) \tag{5.8}$$

$J(W)$ is the discrimination power of a given data projection W. The higher discrimination power is, the better separable will be. Let define the covariance matrix of projected features vector, as written below.

$$
\begin{aligned}
S_w^{proj} &= \sum_{k=1}^{L} E\left[(y^k - E[y^k])(y^k - E[y^k])^T\right] \\
&= \sum_{k=1}^{L} E\left[W^T(x^k - E[x^k])(x^k - E[x^k])^T W\right] \\
&= W^T \left\{ \sum_{k=1}^{L} E\left[(x^k - E[x^k])(x^k - E[x^k])^T\right] \right\} W \\
&= W^T S_w W = [\lambda]
\end{aligned} \tag{5.9}
$$

From Eq. (4.2), Eq.(4.3), Eq. (5.8), and Eq. (5.9), the discrimination power of both CPCA-based and APCA-based projected features vector can be compared as below.

$$
J_{CPCA}(W) \quad = trace(C_g^{proj}) = \textstyle\sum_{i=1}^{m} \lambda_i^{CPCA} \tag{5.10}
$$

and

$$
\begin{aligned}
J_{APCA}(W) &= trace\left((S_w^{porj})^{-1} S_b\right) = trace\left((S_w^{porj})^{-1}\right) \\
&= \sum_{i=1}^{m} \frac{1}{\lambda_i^{APCA}}
\end{aligned} \tag{5.11}
$$

where $S_b = I$ (not considering S_b). As described previously, one of the eigenvalues of global covariance or within class scatter is close to zero, it makes the discrimination power of APCA projected features vector to be very large. Therefore, the Eq. (5.10) and Eq. (5.11) can be compared as below.

$$
J_{CPCA}(W) \ll J_{APCA}(W) \tag{5.12}
$$

Because both of optimum criteria are satisfied, the APCA provides better performance than the CPCA. This fact will be proved by experimental data in the next sections.

5.3.3 Predictive DLDA

The face recognition using the DCT-based holistic features of face images and one-dimensional DLDA has good and stable performance for both small and large

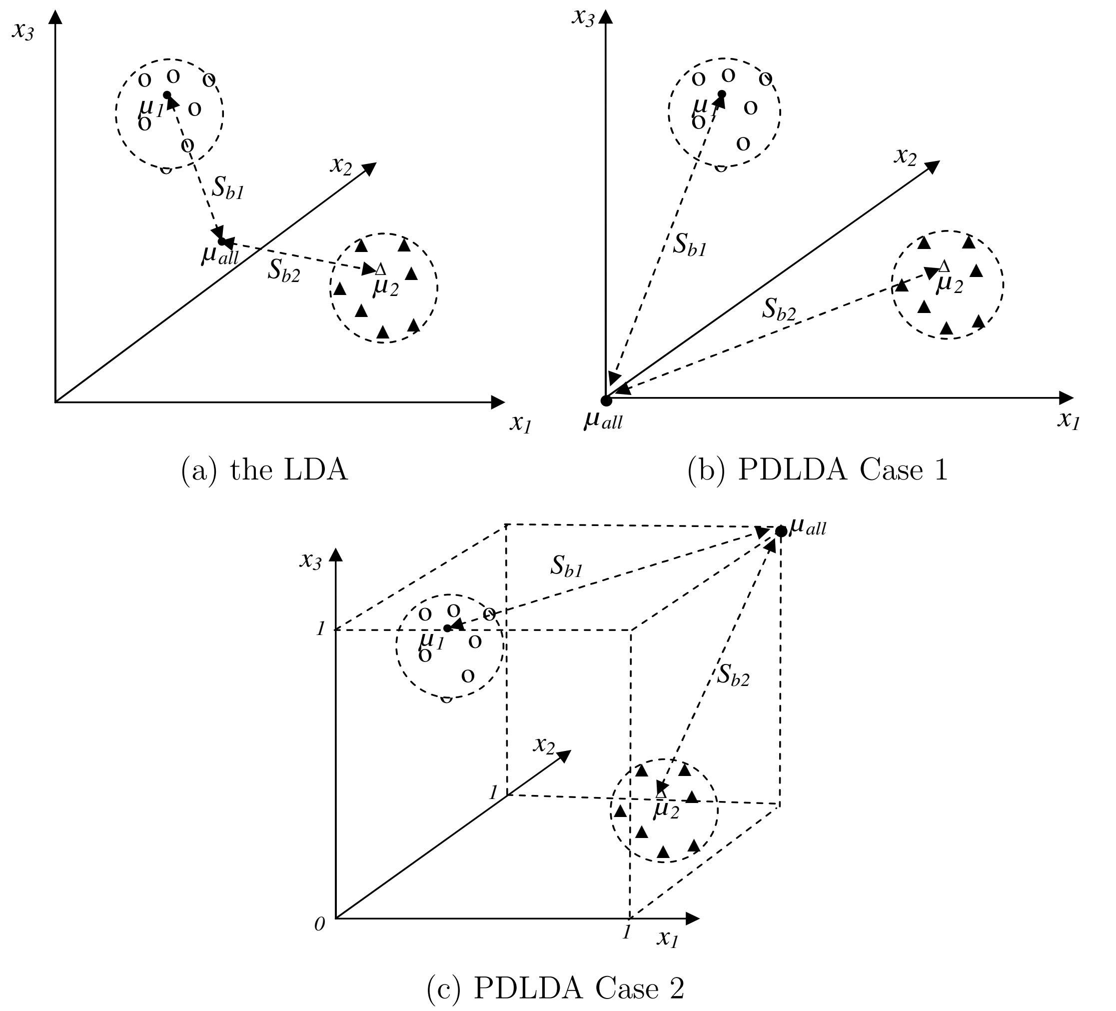


FIGURE 5.3: The illustration two classes clustering of three-dimensional data.

sample size and shown in Ref. [33, 35, 38]. However, it has to retrain all data samples to obtain optimum projection matrix when new data samples enter into the system. It has to be done because the S_b depends on the global means, which has to be recalculated when new data sample comes. In order to avoid this problem and to decrease the computation load of the DLDA algorithm, we develop a predictive LDA (PDLDA), which is derived by redefining the global mean μ_a as a constant value for all samples.

Suppose, we have the data cluster of two classes in three-dimensional which is normalized in the range [0-1], shown in Fig. 5.3(a). Using the Eq. (4.7), the S_b of this case can be written as bellow.

$$S_b^{org} = P(x^1)(\mu_1 - \mu_a)(\mu_1 - \mu_a)^T + P(x^2)(\mu_2 - \mu_a)(\mu_2 - \mu_a)^T \tag{5.13}$$

Based on the LDA criterion as written in the Eq. (4.5), more class separable can be achieved by maximize the determinant of $W^T S_b W$ and minimize the determinant $W^T S_w W$ as much as possible. Another strategy that can be applied for satisfying that criterion without recalculating the S_b is to redefine the constant global mean using three cases model below. Each model should make the S_b have the same basic structure as the original one S_b^{org}: it has positive determinant and symmetry matrix.

1. If the μ_a is moved in the origin point (μ_a is equal to null vector, $\mu_a = [000]^T$), as shown in Fig. 5.3(b)), the S_b of this case can be simplified as

$$S_b^1 = P(x^1)(\mu_1 \mu_1^T) + P(x^2)(\mu_2 \mu_2^T) \tag{5.14}$$

   In this model, the S_b^1 has less computational complexity, does not require to recalculate global mean, and has the same characteristic as the original one in term of symmetric matrix and separability. the separability means that $|W^T S_b^1 W|$ is probably higher than the $|W^T S_b^{org} W|$ which make the data cluster be more separable than that of the original one. This condition make the discrimination power of the data projection is higher than that of the original the data projection.

2. If the μ_a is moved to the maximum value of the range (μ_a is equal to one vector, $\mu_a = [1\ 1\ 1]^T$, as shown in Fig. 5.3(c)), the S_b of this case can be written as

$$S_b^2 = P(x^1)(\mu_1 - [1])(\mu_1 - [1])^T + P(x^2)(\mu_2 - [1])(\mu_2 - [1])^T \tag{5.15}$$

   Event though the predictive S_b^2 has the same computational complexity as in the Eq.(5.13) but it does not require to recalculate global mean, and it has the same characteristic as the original one: symmetric matrix and probably higher discrimination power data projection, as described previously.

3. If the μ_a is forecasted by calculating it from l sub-sample data which is randomly selected from L data samples (i.e $l \ll L$), the S_b of this case can be written as:

$$\begin{aligned} S_b^3 = P(x^1)(\mu_1 - \mu_p)(\mu_1 - \mu_p)^T + \\ P(x^2)(\mu_2 - \mu_p)(\mu_2 - \mu_p)^T \end{aligned} \tag{5.16}$$

   Statistically, the forecasted μ_a (μ_p) has close event equal value to the original μ_a which make the predictive S_b be $S_b^3 \cong S_b^{org}$. Therefore, It absolutely make the discrimination power of the data projection be the almost the same as that of the original once with the same computational complexity as that of the second model.

From these cases, by setting up the predictive global mean as a constant vector, make first case have the smallest computational complexity of that of the others.

For n-dimensional data, all of the predictive S_b can be generalized as the following equation.

$$S_b^p = \sum_{k=1}^{L} P(x^k)(\mu_k - \mu_p)(\mu_k - \mu_p)^T \tag{5.17}$$

Furthermore, because the μ_p is constant, the S_b can be updated using of the following equation when a new class, x^{new}, comes into the system.

$$\begin{aligned} S_b^p &= \sum_{k=1}^{L} P(x^k)(\mu_k - \mu_p)(\mu_k - \mu_p)^T + \\ &\quad P(x^{new})(\mu_{new} - \mu_p)(\mu_{new} - \mu_p)^T \\ &= S_b^{old} + S_b^{new} \end{aligned} \tag{5.18}$$

By substituting the S_b with the S_b^p of CLDA eigen analysis then we get the optimum projection matrix called as PDLDA projection matrix (W_{PDLDA}). The optimum W_{PDLDA} is constructed by selecting small number (m) of eigenvectors which correspond to the large eigenvalues. By using the W_{PDLDA}, the projected features of the both training and querying data set can be performed as the LDA has done using:

$$Y_i^k = W_{PDLDA}^T X_i^k \tag{5.19}$$

Then, the Euclidean distance based on nearest neighbour rule is implemented for classification.

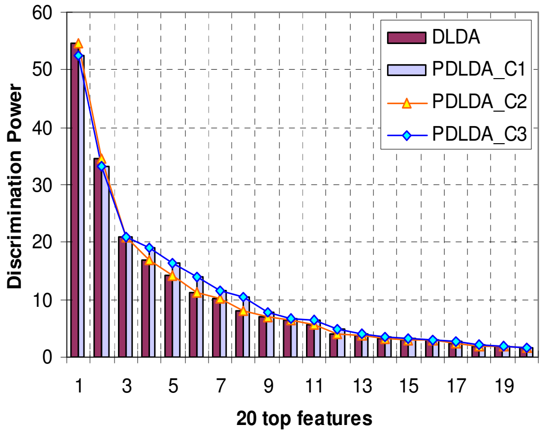
In order to know the strength of the our proposed method to perform feature clustering, we determine the discrimination power (DP) of the proposed algorithms and compare them with that of DLDA, CPCA, and APCA. The DP was determined using the following procedures in well-known ORL database.

1. Determine the projected data (Y_i^k) using the Eq. (5.19),
2. Determine the within and between class scatter of projected data (Y_i^k) called as S_w^{Pro} and S_b^{Pro} as the same as done by LDA respectively, and
3. Calculate the DP of the projected data can be done by substituting S_w and S_b of Eq. (5.8) with S_w^{Pro} and S_b^{Pro} respectively.

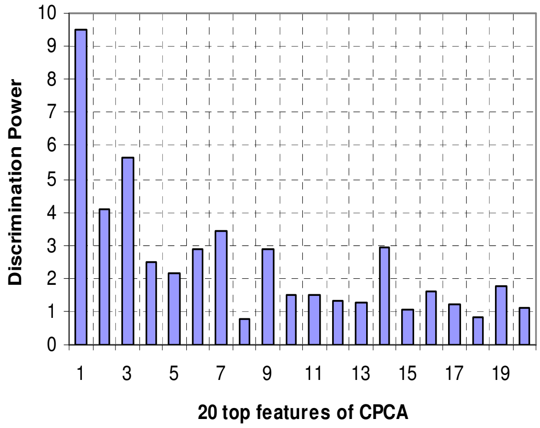
The result shows that the PDLDA has closely the same DP as the DLDA algorithm and higher DP than both of CPCA and APCA, as shown in Fig. 5.4. It means that the almost all of the classification information of face image is place in few top discriminant vectors of PDLDA as the same as that of the DLDA. However, the CPCA and APCA classification information spread to all over features vector therefore the CPCA and APCA will require more eigenvectors to get the optimum projection than DLDA and PDLDA. Between the CPCA and APCA, the second one has higher value of DP even though it has the same spreading shape of the classification information. It proves that the APCA has satisfied the optimum criterion (Eq. (5.12)).

Regarding to time complexity of recalculating S_b using Eq. (5.18), it requires: $(d+d^2)$ multiplication operations for case 1, $(d+d^2)$ multiplication and d addition operations for case 2 and 3. However, the original one using Eq. (4.7) requires $(L+1)(d+d^2)$ multiplication and $(L+1)d$ addition operations, where $L+1$ is total class member of data training and d is the dimensional size of vector μ_i, μ_a, or μ_p.

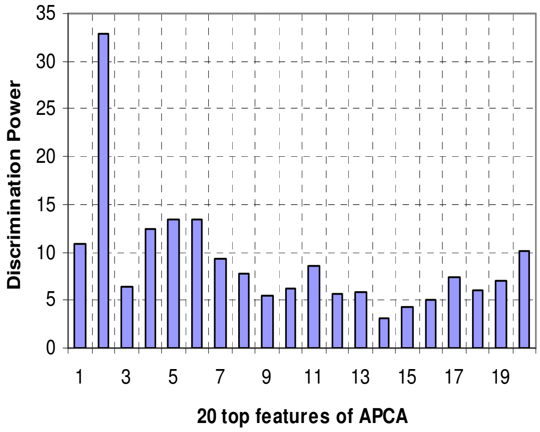
By using these predictive S_b^p to substitute the S_b in Eq.(4.5), the optimum projection matrix W can be obtained by solving the eigen problem of matrix $((S_w^{-1}S_b^p)$ using the direct-LDA (DLDA) algorithms. During the training, the projected data training can be given by Eq. (4.4) as done by the PCA and LDA methods. For



(a) DLDA and three cases of PDLDA



(b) Classical PCA



(c) Alternative PCA

FIGURE 5.4: The Discrimination power of our proposed methods compared to established PCA and LDA

recognition, a Euclidean distance based on nearest neighbour rule is implemented for classification.

5.3.4 Multi-Stage Classifier

In order to get better classification performance, we will apply multi-stage classifiers, i.e. PDLDA as the first stage classifier and the PDLDA or APCA as the second stage classifier. The function of the first stage classifier is to determine several class candidates which are closely to the query image and then the second stage classifier determines the best similarity of query image to the class candidates. The second stage classifier will determine strongly the face likeness because it just compares the face image querying to around class candidates. However, the multi-stage requires longer time processing than that of single stage methods.

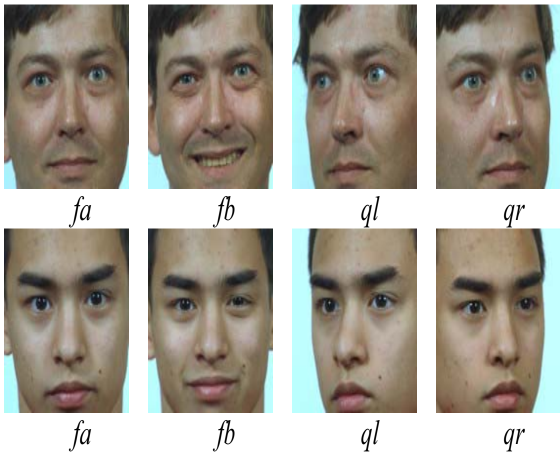


FIGURE 5.5: Example of face pose variation of single subject of FERET database.

Other strategy can be used to integrate the APCA and PDLDA by projected the image features into a new project data using both the APCA and PDLDA projection matrix (W_{APCA} and W_{PDLDA}).

$$y_i^k = W_{APCA} W_{PDLDA} x_i^k \tag{5.20}$$

5.4 Experimentals and Results

The experiments were carried out using several challenging face databases: the ORL database [46, 47], YALE database [49], ITS-Lab. Kumamoto University database [33, 35, 38], INDIA database [53], and FERET database [54]. The example of face pose variations of ORL database is shown in Fig. 2.6(d), ITS-Lab database as shown on Fig. 2.6(a), the Indian database as shown on Fig. 2.6(b), the YALE database is shown on Fig. 3.3, and the rest FERET database is shown in Fig. 5.5. From FERET database, we selected about 508 face classes and each class has 4 images (*fa, fb, ql, qr*). From each class, half of class member were selected as training samples and the remaining as querying samples.

All the experiments were performed in those databases with the following assumptions:

1. the face image size is 128 x128 pixels (28 pixels/cm) with representing using 24 bit for color image and 8 bit of for grayscale image per pixel,

Table 5.1: The comparison of the recognition rate of the proposed to established algorithms.

No	Methods	Features Dimensions	Recognition Rate (%) of				
			ORL	YALE	ITS	INDIA	FERET
1	2DPCA	10x128	95.50	87.88	88.62	87.19	82.61
2	$(2D)^2$PCA	10x10	95.50	87.88	88.62	87.19	81.77
3	$(2D)^2$PCALDA	8x8	97.50	**92.12**	96.41	86.63	90.77
4	DCT+2DPCA	8x12	96.75	88.49	88.82	88.32	82.10
5	DCT+$(2D)^2$PCA	8x8	96.25	88.49	87.82	88.51	81.59
6	DCT+CPCA	38	95.25	89.70	94.41	96.61	91.78
7	DCT+APCA	38	**97.75**	91.52	**99.80**	**97.18**	**94.00**
8	2DLDA	8x128	96.25	86.67	97.21	88.14	91.48
9	$(2D)^2$LDA	8x8	97.50	88.49	93.21	68.93	84.59
10	DCT+2DLDA	6x12	95.75	90.30	98.80	90.40	91.66
11	DCT+$(2D)^2$LDA	6x6	97.50	90.91	98.60	88.32	91.24
12	DCT+$(2D)^2$PCALDA	8x8	97.75	91.51	97.00	89.45	89.99
13	DCT+DLDA	24	98.50	**95.15**	**99.47**	97.36	96.94
14	DCT+PDLDA_ C1	24	**98.75**	92.12	**99.52**	97.18	96.88
15	DCT+PDLDA_ C2	24	98.25	**95.15**	**99.52**	**97.55**	**96.94**
16	DCT+PDLDA_ C3	24	98.50	**95.15**	**99.47**	97.36	96.88

2. a half of samples of each class/subject members were selected as training faces and testing faces overlapped with the training faces,

3. The training face is included in the query process because there is possibility that the query faces are almost or exact the same as the training face in the real time system, and

4. the dimension of the global features vector of face image is 53 elements of original size (128^2 pixels).

The first experiment, which was performed in all the mentioned databases, investigates the robustness of our proposed methods over the established methods. From the experimental results, all variations of the PDLDA tend to provide the same recognition rate as the DLDA with the same features dimension requirement and outperform over the established methods in all databases as shown in the Table. 5.1. It means that our proposed methods give robust recognition rate in all tested databases. It can be achieved because all of the PDLDA variations almost have the same discrimination power as the DLDA, see Fig. 5.4. It proves that the S_b of PDLDA has the same structure and optimum criterion as that of the DLDA. From the PCA point of view, the APCA also tend to provide better performance than those of CPCA and its variations. It can be achieved because the APCA's projection matrix is more optimally than that of CPCA. The more optimally projection matrix is, the better features cluster will be. Moreover, APCA comparison result proves the APCA has higher discrimination power than the classical PCA (Eq. (5.12)). Overall, the results also prove that the LDA and its variations have better performance than that of any variations of the CPCA.

From the optimum performances of our proposed methods plotted by cumulative match score (CMS) curve as shown in Fig. 5.6, show that all of the PDLDA methods almost give the same recognition rate as that of DLDA and higher recognition rate than that of established methods. This result explains that all of the PDLDA methods have the similar strength/ability to the DLDA for face recognition with the global compact features as the raw input. Furthermore, all of the PDLDA methods require almost the same projected features dimension as that of the DLDA and less than that of PCA as shown in Tables 5.2. It proves that a few top discriminant vectors of the PDLDA contain most of the classification information of face image features as that of DLDA has done. However, this method will require large time processing when features dimensional is large (using original image as raw input). Therefore, we implement the compact global features to overcome large computational cost problem.

Regarding to optimum recognition rate of the APCA, it shows that the APCA is out of performs over the CPCA. In term of rank 1, the APCA gives higher recognition rate for all of the tested databases as shown in Fig. 5.6. It strongly proves that the APCA transforms the input features such that features clusters are more separable after the transformation than that of the CPCA. In other word, the results prove $J_{CPCA}(W) \ll J_{APCA}(W)$ and match with the mathematical analysis as explained in sub-section 5.3.2.

In order to show that the proposed method required less time processing for training, the next experiment was performed. It was done in FERET face database with face features of 53 elements and the training was performed gradually: firstly, it was trained 208 face classes and then added gradually 20 new face classes to the system until 508 face classes. The experimental results were plotted in Fig. 5.7(a). It shows that all variations of the PDLDA require less training time than that of DLDA. Even though, the training time of the PDLDA increase but its average of increment slope is less than half of that of DLDA by about $12.2\text{x}10^{-5}$ and $29.2\text{x}10^{-5}$ respectively. Furthermore, the larger number of classes are, the larger

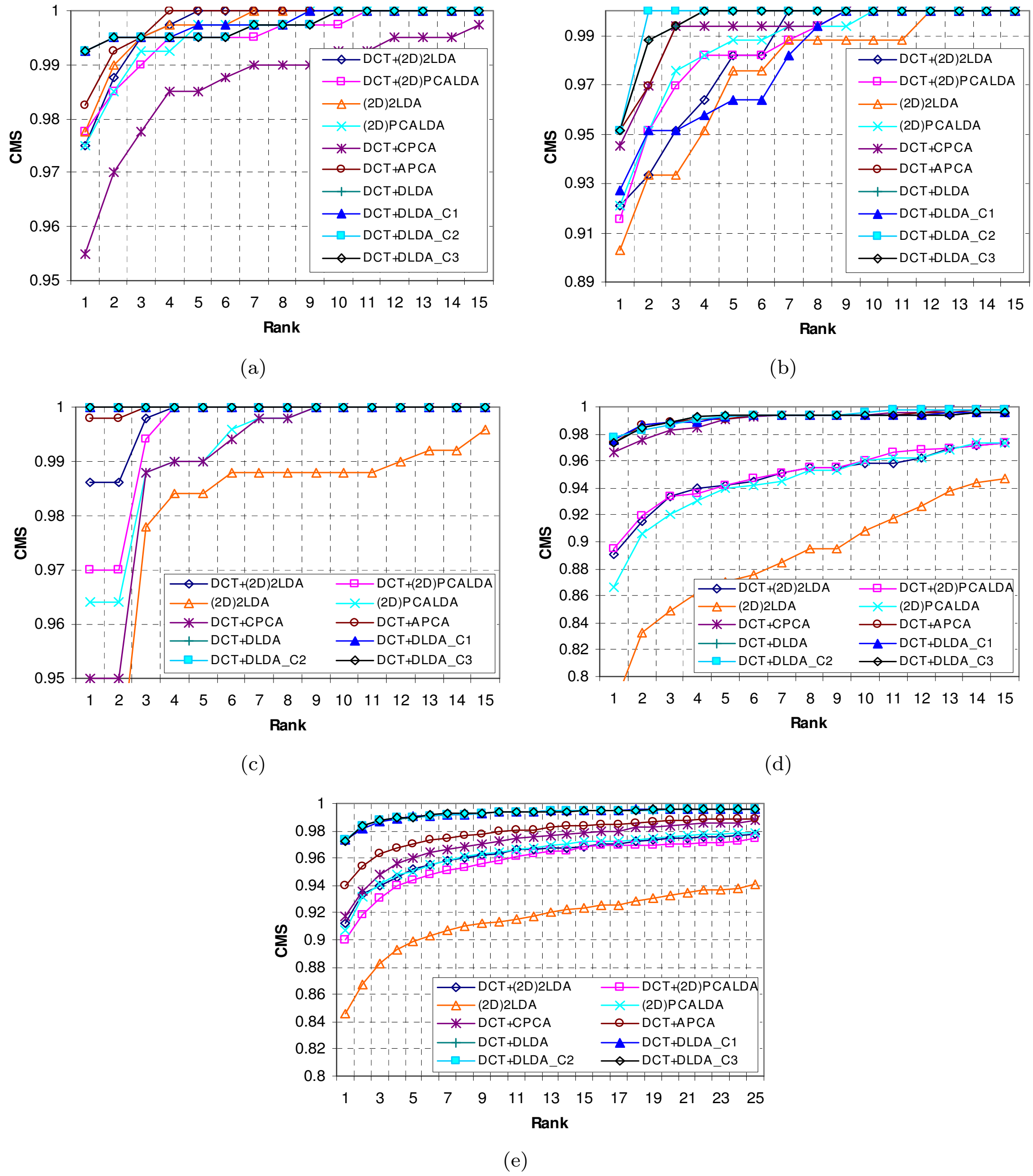


FIGURE 5.6: CMS curve of our proposed methods compared to the established methods on: (a) ORL, (b) YALE, (c) ITS. Lab., (d) INDIA, and (e) FERET database.

increment of training time of the DLDA will be, while the PDLDA almost has constantly training time increment. This result proves that the PDLDA requires less time processing which is one of the strength of the PDLDA, when retraining is done. It can be achieved because the PDLDA has simpler computation complexity than that of DLDA ($d+d^2$ multiplication for PDLDA_C1, $d+d^2$ multiplication and d addition for PDLDA_C2 and for PDLDA_C3, and $(L+1)(d+d^2)$ multiplication and $(L+1)d$ addition for DLDA) as described in section (4.c).

TABLE 5.2: The features dimensional requirement comparison to get the optimum recognition rate.

Methods	Recognition Rate / Dimension				
	ORL	YALE	ITS	INDIA	FERET
DCT+CPCA	30	24	30	38	38
DCT+$(2D)^2$PCA	6x6	12x12	12x12	8x8	8x12
DCT+APCA	30	24	28	40	38
DCT+$(2D)^2$LDA	6x6	10x10	6x6	4x4	6x12
DCT+$(2D)^2$PCALDA	8x8	12x12	8x8	8x8	8x8
DCT+DLDA	28	14	22	22	14
DCT+PDLDA_ C1	32	12	20	36	14
DCT+PDLDA_ C2	30	12	22	22	14
DCT+PDLDA_ C3	30	14	22	22	14

Regarding to the recognition rate stability of our proposed methods as shown in Fig. 5.7(b), they give almost the same stability recognition rate as that of the DLDA. It means that the PDLDA has the same structure as the DLDA but it has simpler computation complexity. It is an alternative algorithm for features cluster of large sample size databases, which requires much retrain processing.

In the next experiment, we investigated the effect of multi-stage classifier to the recognition rate that was done in INDIA database, which represents small size database and FERET database, which represents large size database. In this case, we implemented the PDLDA as the first-stage classifier and the APCA or DLDA as the second stage classifier. The PDLA is chosen because it has less computational cost than that of DLDA and it has robust recognition rate in large size database. Overall, the experimental results show that all variants of multi-stage tend to provide higher performance than the baseline (PDLDA), as shown in Fig. 5.8. For INDIA database (Fig. 5.8(a)), the highest recognition rate improvement is given by the PDLDA_C2 +APCA and DLDA. It means that both of the second-stage classifiers provide the same performance. However, for the FERET database (Fig. 5.8(b)), the highest improvement is given by the PDLDA_ C1 + DLDA. This result explains that the PDLDA_ C1 and PDLDA_ C2 tend to provide better result than the PDLDA_ C3, which prove the W^TS_bW of both the PDLDA_ C1 and PDLDA_ C2 have higher determinant than that of the DLDA (see the illustration in the Fig. 5.3). Based on the optimum criterion of the LDA, the more separable projected data can be achieved by maximizing the determinant of W^TS_bW as much larger as possible. In other side, the S_b of the PDLDA_ C3 probably has much the same value as that of the DLDA because the global mean was forecasted

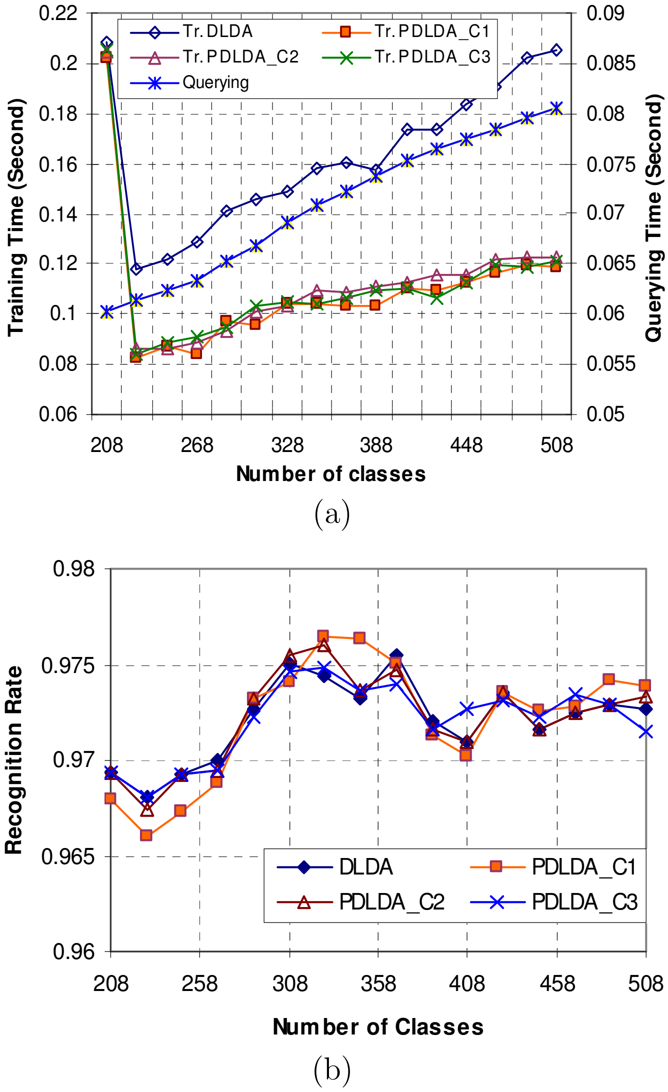
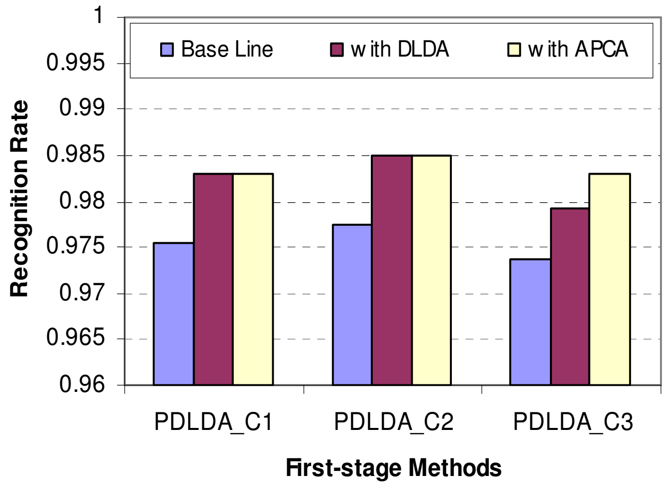


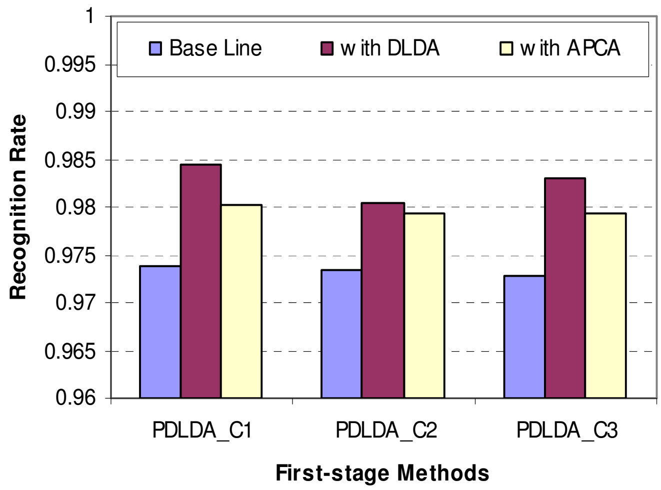
FIGURE 5.7: (a) Training and querying time consumption and (b)The recognition rate stability over the number of classes .

by randomly sampling samples. Statistically, if the N sampling samples is close to the population sample, the forecasted mean will be much the same as the real mean. Therefore, the PDLDA_ C3 combinations give less improvement.

The last test was performed to know the accuracy of our proposed methods compared to the others methods. In this test, we investigate three important parameters, namely false rejection rate (FRR) and false acceptance rate (FAR) and equal error rate (EER). The FAR is the probability of unauthorized user to be falsely as accepted or recognized as legally user. The FRR is the probability legally registered user to be falsely rejected by the system. The system that performs perfect classification or has high accuracy, is denoted by 100% true positive rate and 0%



(a) INDIA databases



(b) FERET databases

FIGURE 5.8: The multi-stage vs. base line recognition rate.

false positive rate or the value of EER is small or close to zero. We use receiver operating characteristics (ROC) as performance evaluation because the ROC is not so needy on the precise variety of test data compared to confusion matrix and it is robust with respect to class skew.

The test was performed on all mentioned databases except on the YALE database because it has the smallest sample size. Those of databases are subjected as predicted positive (known face) and the frontal face image of CVL database is subjected as predicted negative (unknown face). In this case, we add a threshold for the distance measure between features permit rejection of unknown face and verifications of those that are known. In other word, we sent an unknown face and a "claimed" identity to the system for verifications. If the distance between the

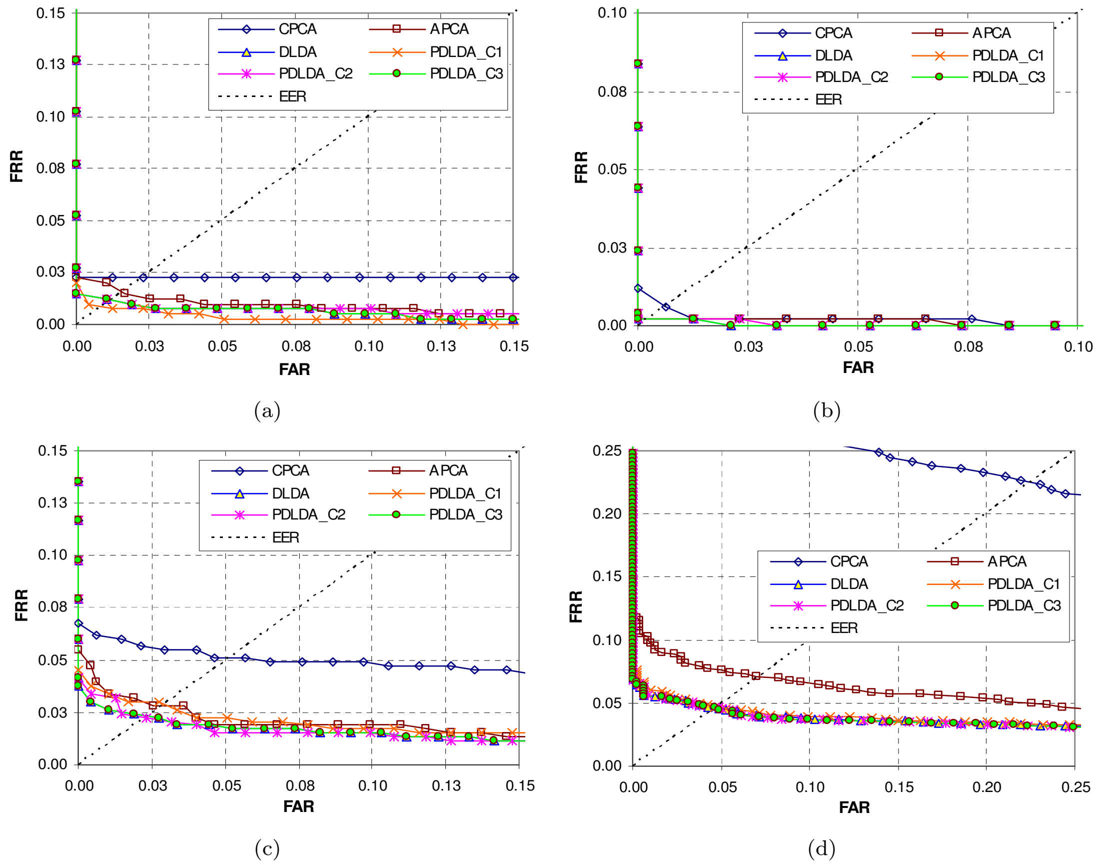


FIGURE 5.9: ROC curve of proposed method compared to the other methods on: (a) ORL, (b) ITS. Lab., (c) INDIA, and (d) FERET database.

face's features to those of database image, which it is being verified less than the given threshold, the claimed, is accepted, otherwise, it is rejected.

The experimental results are plotted using ROC as shown in Fig. 5.9. The results show that the APCA-based approach performs better than the CPCA methods that are denoted by smaller ERR of the APCA than that of the CPCA. In other side, the PDLDA-based methods give much the same performance as the DLDA, which is denoted by the almost the same ERR, but it is better than that of the CPCA and APCA methods. For ORL. database the smallest EER is given by the PDLDA_C1 (0.0127) with the optimum recognition rate 99.25%, for ITS-Lab. database the smallest EER is given by the PDLDA_C1, PDLDA_C2, PDLDA_C3, and DLDA (0.0019) with the optimum recognition rate 100%, for INDIA database the smallest EER is given by PDLDA_C2 (0.0232) with the optimum recognition rate 97.74%, and the rest for FERET database the smallest EER is given by PDLDA_C1, PDLDA_C2, PDLDA_C3, and DLDA (0.0464) with the optimum recognition rate 97.39%. These results support our previous experimental results that the PDLDA-based algorithms have the same powerful as the established DLDA. However, the PDLDA-based algorithms requires less time processing for retraining and the combination of compact global features and the PDLDA gives much higher improvement than recent 2DLDA and two directional 2DLDA. It can be achieved because the compact global features (dominant frequency content + moment information) already contain most discriminant information denoted by high enough discriminatory power as shown in Fig. 5.1.

Finally, the real-time test was done on ITS Lab. face database lab with consist of 98 subjects. In this test, we probed non-training subject into the system as querying image in order to know the false acceptance rate (FAR). When the non-training subject was not recognized then it was registered as the new training subject and was continued to the querying process. In this test, we wanted to know the false rejection rate (FRR), and the time processing of retraining and querying. From 50 subjects that were tested, this process provides FRR : 2% and the FAR: 4% with short time processing for retraining and querying by about 0.6437 and 0.0235 seconds respectively. The real-time was test on PC with specification : Core Duo 1.60 GHz processor and 2 GB RAM. This result can be achieved because the PDLDA has the same characteristics as that of the DLDA in term of the discriminant power and the retraining does not require recalculating the S_b, which depends on the global mean.

5.5 Conclusion and Future Works

From all of the experimental results have been prove our hypothesis as follows. Firstly, the global features as a dimensional reduction of raw image has been successfully implemented with good achievement when the one-dimensional APCA, DLDA, and DLDA as features clustering. Secondly, PDLDA-base face recognitions has been proved that they require less time processing for training and re training. Thirdly, the optimum recognition performance can be achieved by integrating the PDLDA and APCA or DLDA which the almost the same time processing. Fourthly, the proposed method (APCA and PDLDA) outperforms over the recent sub-space methods (2DPCA and 2DLDA based methods) and the PDLDA can used as alternative solution to avoid recalculating the S_b and global mean of the retraining process of DLDA. In addition, the PDLDA is alternative solution for large size data clustering because it does not depend on the global means. The rest, the real-time PDLDA based face recognition provides good enough achievement in term of recognition rate, FRR and FAR with short time processing for retraining and querying. However, the real-time experiments just performed in small subjects, we plan to test in large subjects in order to know its robust performance.

In order to get more precise verification result, we will consider more local features analysis involving eyes, nose, mouth, and context information of the face image.

Acknowledgement

I would like to send my appreciation and my great thank to all people who participate in building ITS-Lab face database and to the owner of ORL, INDIA, and FERET face database.

Chapter 6

Conclusion and Future Work

In this research we propose some alternative methods to face recognition, which is mainly based on global/holistic features of face image that is extracted using frequency analysis, shape analysis, and skin tone information. Our proposed method has been successfully implemented and from all of the experimental results some conclusions can be written as follows:

1. The dominant frequency features based face recognition method gave good performance achievement which can reduce significantly training time of PCA training time, need less time querying, and requires small facial features size than that of PCA features.

2. Multi-pose face image recognition based on hybrid dominant frequency features has been successfully implemented which give robust performance to face pose variations and requires low computational cost, requires little memory space, and can overcome retraining problem than that of PCA or LDA based face recognition methods.

3. Frequency analysis is an efficient way to create the compact features which is good concept of reducing memory space requirement and computational load of LDA and PCA. Between the DCT and the DWT frequency analysis, the first one performed better face image representation that was shown by higher success rate and smaller EER than DWT's success rate and EER.

4. The global features tends to provide good performance achievement when the one-dimensional APCA, DLDA, and PDLDA as features clustering which

requires less time processing for training and re training. Lastly, the integration of APCA and PDLDA outperforms over the recent sub-space methods (2DPCA and 2DLDA based methods) and the PDLDA can used as alternative solution to avoid recalculating the S_b and global mean of the retraining process of DLDA. In addition, the PDLDA is alternative solution for large size data clustering because it does not depend on the global means.

5. The PDLDA classifier with global features as the input data is recommended to be applied for incremental data face recognition, which requires retraining process for both off-line and real-time system.

In the future this works will continue to more precise verification result by considering more local features analysis involving eyes, nose, mouth, and context information of the face image. Because our proposed method requires short time processing, it will be possible to be implemented as the real-time face security system, such as attendance register system.

Appendix A

Interface of Face Recognition

The interface of face recognition was created using C++ builder version 6.0 and OpenCv library which is open and free library from Intel cooperation. The face recognition interface mainly consist of some interfaces: initial, face detection, face register, face training, and face recognition.

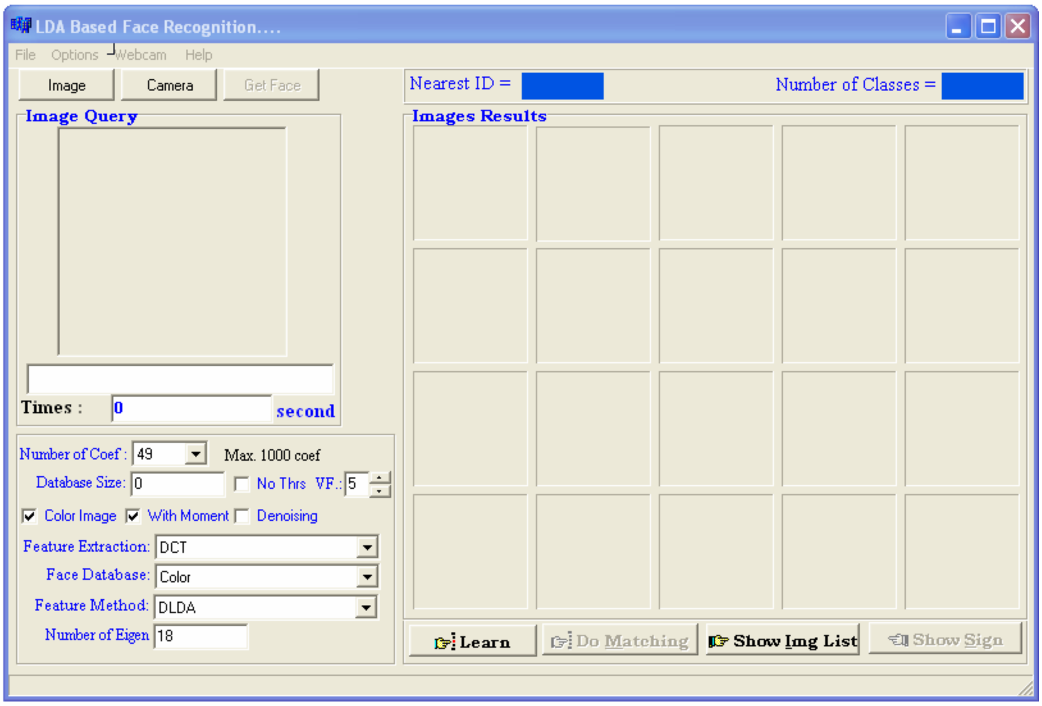


FIGURE A.1: Initial interface of face recognition system

The initial face recognition interface can be seen in Fig. A.1. Using this initial interface, the user can setup the initial parameters that is required by the face recognition system including features vector size, number of eigen vectors, verification threshold, and face database. The default value of those parameters for: features vector size is 49, number of eigen vectors is 18, verification threshold is 5.

Figure A.2: Face Detection Interface

For face detection, we implemented haar-like based face detection which was offered by the OpenCv. This face detection has been proven that it could detect robustly human face both in real-time and off line. The interface of face detection can be seen in Fig. A.2. To operate the face detection, the user has to follow these procedures: click the camera button in order to show the web camera interface and then click the get face button in order to detect the face. If the face is detected successfully, the detected face will be shown in the *Image Query* container.

Face registration interface was designed to enrol the face images of into the database collection and to perform the retraining process as shown in Fig A.3. The face registering procedure: firstly, several face variations of a subject at least 6 face images are taken by click the face button and then the detected face move into *Facec List* container, and then click the *retraining* button. The retraining process recalculates the optimum projection matrix and projected input features. In this case, the face image features, optimum projection matrix, S_b, and S_w are saved for next process.

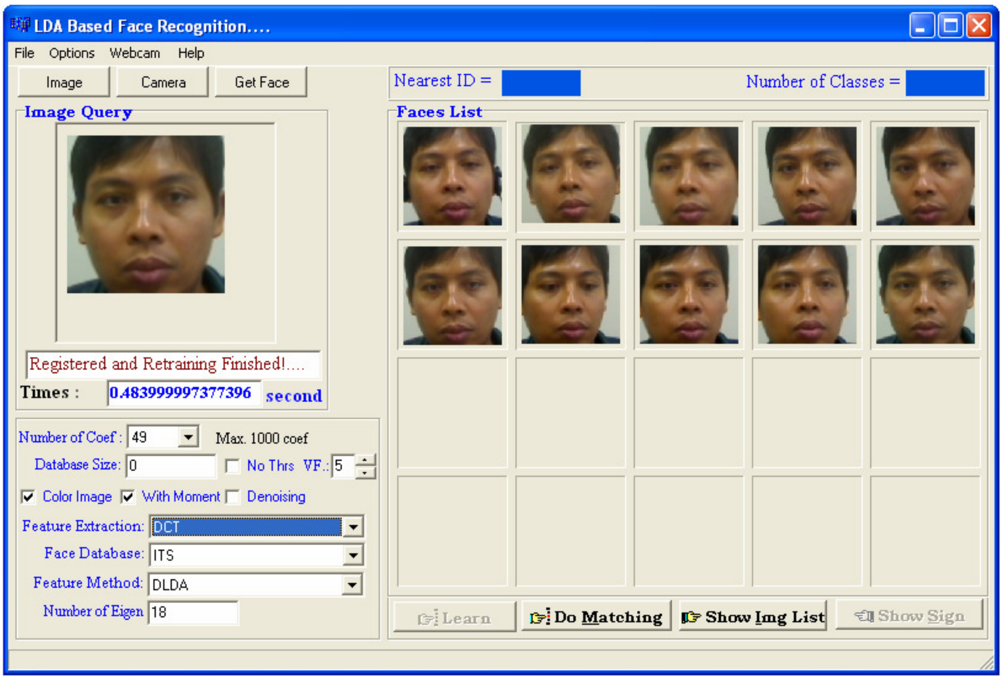


FIGURE A.3: Face registration interface.

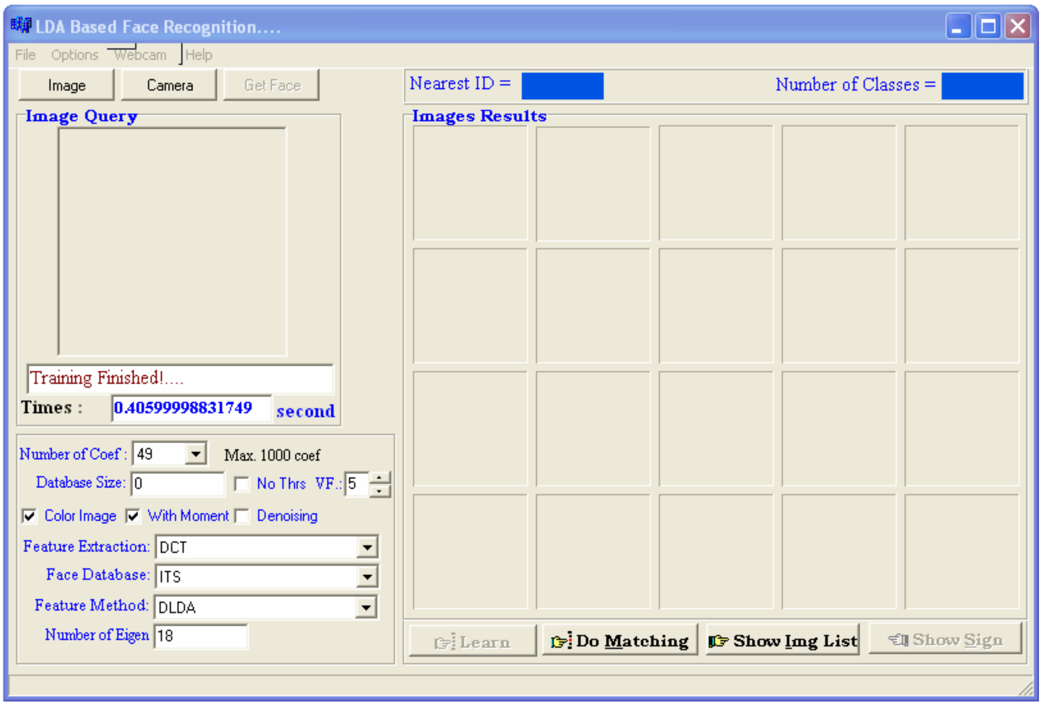


FIGURE A.4: Training Interface

Training interface is used to defined the optimum projection matrix in order to get high power discriminant of the projected data by clicking the Training button in Fig A.4. Before performing the training process the user has to set the initial paremeters as metioned previously.

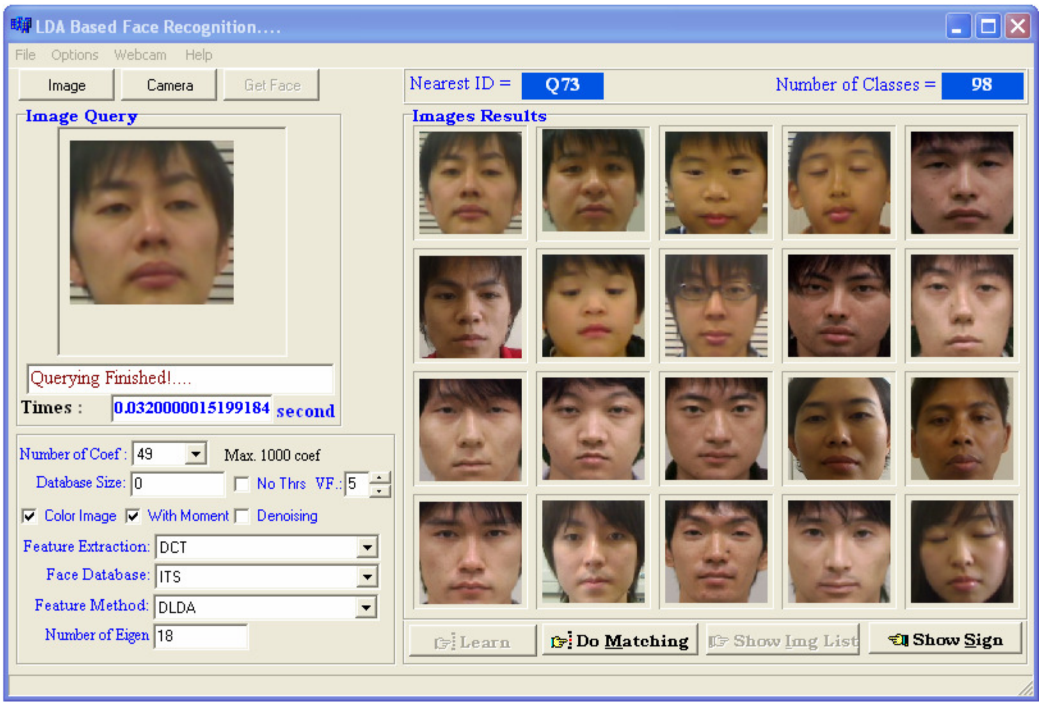


FIGURE A.5: Querying Interface

Recognition interface is an GUI tool to demonstrate the recognition ability of the proposed system. The input of this interface can be image or realtime video. For example, if we input the query image as shown in Fig A.5 then click the Matching button, the system will give the 20 face images that mostly similar with the input. The best likenes face image with the input face is the first image in the *images result* container or top left face image which corresponds to the nearest ID=Q73.

Bibliography

[1] http://biometrics.cse.msu.edu/Presentations/

[2] W. Zhao, R. Chellappa, and A. Rosenfeld, "Face Recognition: A Literature Survey", ACM Computing Surveys, 35, pp.399-458, December 2003.

[3] X. Zhang and Y. Gao,"Face Recognition Across Pose: A Review", Pattern Recognition, April 2009, doi:10.1016/j.patcog.2009.04.017.

[4] W. Gao, B. Chao, S. Shan, X. Chen, D. Zhou, X. Zhang, and D. Zhao, "The CAS-PEAL Large-Scale Chinese Face Database and Baseline Evaluations", IEEE Transactions on System, Man, and Cybernetics-Part A: Systems and Human, vol. 38, no. 1, pp. 149-161, January 2008.

[5] Z. M. Hafed and M. D. Levine, "Face Recognition Using the Discrete Cosine Transforms", International Journal of Computer Vision, vol. 43, no. 3, pp. 167-188, 2001.

[6] M. Sharkas, "Application of DCT Blocks with Principal Component Analysis for Face Recognition", Proc. the 5th WSEAS Int. Conf. on Signal, Speech, and Image Processing, pp. 107-111, Corfu, Greece, 2005.

[7] C. Garcia, G. Zikos, and G. Tziritas, "Wavelet Packet Analysis for Face Recognition", Image and Vision Computing, vol. 18, pp.289-297, 2000.

[8] D-Q. Dai and P. C. Yuen, "Wavelet Based Discriminant Analysis for Face Recognition", Applied Mathematics and Computation, vol. 175, pp. 307-318, 2006.

[9] M. Yoshida, T. Kamio, and H. Asai, "Face Image Recognition by 2-Dimensional Discrete Walsh Transform and Multi-Layer Neural Network", IEICE Trans. Fundamentals, vol. E-86A, pp. 2623-2627, October 2003.

[10] S.Z. Li and A. K. Jain, "Handbook of Face Recognition", Springer New York USA, 2005.

[11] G.C Feng, P.C. Yuen, and D.Q. Dai, "Human Face Recognition Using PCA on Wavelet Sub-Band", Journal of Electronic Imaging, vol. 9, no. 2, pp. 226-233, 2000.

[12] F.J. Huang, Z. Zhou, H-J. Zhang, and T. Chen, "Pose Invariant Recognition", Proc. the 4th IEEE International Conference on Automatic Face and Gesture Recognition, pp. 245-250, Grenoble, France, 2000.

[13] C.E. Jacobs, A. Finkelstein, and D.H. Salesin, "Fast Multiresolution Image Querying", Computer Graphics Annual Conference Series, vol. 29, pp. 277-286, 1995.

[14] E. Horst, "Kalman Filtering for Pose-Invariant Face Recognition", Proc. ICIP 2006 International Conference on Image Processing, pp. 2037-2040, Atlanta, USA, 2006.

[15] T. Satonaka and K. Uchimura, "An Adaptive Metric Learning Procedure for Reconfigurable Facial Signature Authentication", International Journal of Image and Graphics World Scientific Publishing Company, pp. 597-606, 2005.

[16] N. Shams, I. Hosseini, M.S. Sadri, and E. Azarnasab, "Low Cost FPGA-Based Highly Accurate Face Recognition System Using Combined Wavelets with Subspace Methods", Proc. ICIP 2006 International Conference on Image Processing, pp. 2077-2080, Atlanta, USA, 2006.

[17] P. Shih and C. Liu, "Improving The Face Recognition Grand Challenge Baseline Performance Using Color Configurations Across Color Space", Proc. ICIP 2006 International Conference on Image Processing, pp. 1001-1004, Atlanta, USA, 2006.

[18] M. Turk and A. Pentland, "Eigenfaces for Recognition", Journal of Cognitive Neuroscience, vol. 3, no. 1, pp. 71-86, 1991.

[19] J. Yang, D. Zhang, A. F. Frange, and J-Y. Yang, Two Dimensional PCA: A New Approach to Appearance-Based Face Representation and Recognition", IEEE Transaction on Pattern Analysis and Machine Intelligence, vol. 26, no. 1, pp. 131-137, January 2004.

[20] Y-G. Kim, Y-J. Song, U-D. Chang, D-W, Kim, T-S. Yun, and J-H. Ahn, "Face Recognition Using a Fusion Method Based on Bidirectional 2DPCA", Applied Mathematics and Computation, vol. 205, pp. 601-607, 2008.

[21] D. Zhang and Z-H. Zhou, "$(2D)^2 PCA$: Two-Directional Two-Dimensional PCA for Efficient Face Representation and Recognition", Neurocomputing, vol. 69, pp. 224-231, 2005.

[22] J. Meng and W. Zhang, "Volume Measure in 2DPCA-Based Face Recognition", Pattern Recognition Letter, vol. 28, pp. 1203-108, 2006.

[23] K. Etemad and R. Chellappa, "Discriminant Analysis for Recognition of Human Face Images", J. opt. Soc. Am. vol. 14, no. 8, pp. 1724-1733, 1997.

[24] H. Yu and J. Yang, "A Direct LDA algorithm for High-Dimensional Data-with Application to Face Recognition", Pattern Recognition, vol. 34, pp. 2067-2070, 2001.

[25] J. Wang, W. Yang, Y. Lin, and J. Yang, "Two-Directional Maximum Scatter Difference Discriminant Analysis for Face Recognition", Neurocomputing, vol. 72, pp. 352-358, 2008.

[26] S. Noushath, G. H. Kumar, and P. Shivakumara, "$(2D)^2 LDA$: An Efficient Approach for Face Recognition", Pattern Recognition, vol. 39, pp. 1396-1400, 2006.

[27] J. Lu, K.N. Plataniotis, and A.N. Venetsanopoulus, "Face Recognition Using LDA-Based Algorithm", IEEE Transactions on Neural Networks, vol. 14, no. 1, pp. 195-200, 2003.

[28] W. Chen, J-E. Meng, and S. Wu, "PCA and LDA in DCT Domain", Pattern Recognition Letter, vol. 3, no. 26, pp. 2474-2482, 2005.

[29] W-S. Zheng, J. H. Lai, and S. Z. Li, "1D-LDA vs. 2D-DLDA: When is Vector-Based Linear Discriminant Analysis Better Than Matrix-Based?", Pattern Recognition, vol. 41, pp. 2156-2172, 2008.

[30] IGPS. Wijaya, K. Uchimura, and Z. Hu, "Why the Alternative PCA Provide Better Performance for Face Recognition", The International Workshop on Image Analysis for Multimedia Interactive Services (WIAMIS), pp. 169-172, London, UK, May 2009.

[31] IGPS. Wijaya, K. Uchimura, and Z. Hu, "Pose Invariant Color Face Recognition Based on Frequency Analysis and DLDA with Weight Score Classification", The World Congress on Computer Science and Information Engineering (CSIE 2009), Los Angeles USA, March 2009 (CD-ROM).

[32] IGPS. Wijaya, K. Uchimura, and Z. Hu, "Face Recognition Based On Dominant Frequency Features And Multiresolution Metric", The 2nd International Conference on Innovative Computing, Information and Control (ICICIC), Kumamoto, Japan, September 2007 (CD-ROM).

[33] IGPS. Wijaya, K. Uchimura, and Z. Hu, "Face Recognition Based on Dominant Frequency Features and Multiresolution Metric", International Journal of Innovative Computing, Information and Control (IJICIC), vol. 5, no. 3, pp. 641-651, March 2009.

[34] IGPS. Wijaya, K. Uchimura, and Z. Hu, "Face Recognition Based on Holistic Information and Minimum Mahalanobis Classifier", ACCV sub-space workshop 2007, pp. 53-60, Tokyo, Japan, November 2007.

[35] IGPS. Wijaya, K. Uchimura, and Z. Hu, "Pose Invariant Face Recognition Based on Hybrid Dominant Frequency Features", IEICE Transactions on Information and Systems, vol. 91-D, no. 8, pp. 2153-2162, 2008.

[36] IGPS. Wijaya, K. Uchimura, and Z. Hu, "Multipose Face Recognition Based on Frequency Analysis and Modified LDA", The Workshop of Image Electronics and Visual Computing 2007, Cairns Australia, November 2007 (CD-ROM).

[37] IGPS. Wijaya, K. Uchimura, and Z. Hu, "Modified LDA Classifier in Multi resolution Wavelet Domain for Multipose Face Recognition", Sixth International Conference on Electrical Engineering, Cairo, Egypt, May 2008 (CD-ROM).

[38] IGPS. Wijaya, K. Uchimura, and Z. Hu, "Multipose Face Recognition Based on Frequency Analysis and Modified LDA", The journal of the IIEEJ, vol. 37, no. 3, pp. 231-243, 2008.

[39] IGPS. Wijaya, K. Uchimura, and Z. Hu, "Improving the Pose Invariant Face Recognition Using YCbCr Color Configuration", ITE Annual convention 2008, Fokuoka Japan, September 2008 (CD-ROM).

[40] IGPS. Wijaya, K. Uchimura, and Z. Hu, "Improving the Pose Invariant Face Recognition Using Double Frequency Analysis and Color Information", The International Session of 2008 joint Conference of IEEE Computer Society Fukuoka Chapter, Oita Japan, September 2008 (CD-ROM).

[41] IGPS. Wijaya, K. Uchimura, and Z. Hu, "Improving the Pose Invariant Face Recognition Performances Using Multi-Stage Classifier", The Proc. of the Forum on Information Technology (FIT 2009), Sendai Japan, September 2009 (CD-ROM).

[42] IGPS. Wijaya, K. Uchimura, and Z. Hu, "Improving the PDLDA based face recognition using lighting compensation", The Workshop of Image Electronics and Visual Computing 2010, Nice France, March 2010 (CD-ROM).

[43] IGPS. Wijaya, K. Uchimura, and Z. Hu, "Real-Time Predictive LDA base Face Recognition", The 2010 IEICE General Conference, Sendai Japan, March 2010 (CD-ROM).

[44] Z. Zhang and Y. Zhao, "Multiple Description Image Coding Based On Fractal", International Journal of Innovative Computing, Information and Control, vol. 3, no. 6(B), pp. 1615-1623, 2007.

[45] http://fourier.eng.hmc.edu/e161/lectures/dct/node3.html.

[46] F. Samaria and A. Harter, "Parametrization of a stochastic model for human face identification". The 2nd IEEE Workshop on Applications of Computer Vision, pp. 138-142, Sarasota Florida, 1994.

[47] http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html.

[48] R. C. Gonzalez and R. E. Woods, Digital Image Processing, Pearson Prentice Hall, USA, Third edition, 2009.

[49] http://cvc.yale.edu/projects/yalesfaces/yalefaces.

[50] http://homepage.cs.uri.edu/faculty/hamel/pubs/hamel-roc.pdf.

[51] H. Gao and W.D. James, "Why direct LDA is not equivalent to LDA", Pattern Recognition Letter, vol. 39, pp. 1002-1006, 2005.

[52] R.L. Hsu, M. Abdel-Monttaleb, and A.K. Jain, "Face Detection in Color Images", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 24, no. 5, pp. 696-705, 2002.

[53] http://vis-ww.cs.umass.edu/∼vidit/IndianFaceDatabase.

[54] P. J. Philips, H. Moon, S. A. Risvi, and P. J. Rauss, "The FERET Evaluation Methodology for Face Recognition Algorithms", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 22, no. 10, pp. 1090-1104 2000.

[55] Z. Li, D. Lin, and X. Tang, "Non Parametric Discriminant Analysis for face Recognition", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 31, no. 4, pp. 755-761, 2009.