

## DOCTORAL THESIS

### Discriminant analysis algorithms for face recognition

Huang, Jian

*Date of Award:*  
2006

[Link to publication](#)

#### General rights

Copyright and intellectual property rights for the publications made accessible in HKBU Scholars are retained by the authors and/or other copyright owners. In addition to the restrictions prescribed by the Copyright Ordinance of Hong Kong, all users and readers must also observe the following terms of use:

- Users may download and print one copy of any publication from HKBU Scholars for the purpose of private study or research
- Users cannot further distribute the material or use it for any profit-making activity or commercial gain
- To share publications in HKBU Scholars with others, users are welcome to freely distribute the permanent URL assigned to the publication

# **Discriminant Analysis Algorithms for Face Recognition**

**HUANG Jian**

A thesis submitted in partial fulfillment of the requirements  
for the degree of  
Doctor of Philosophy

Principal Supervisor: Prof. YUEN Pong chi

Hong Kong Baptist University

February 2006

# Abstract

Face Recognition research started in the late 70s and a number of algorithms/systems have been developed in the last decade. Among various algorithms, appearance-based approach is one of the promising approaches in face recognition. Linear Discriminant Analysis (LDA) and Principal Component Analysis (PCA) are two most popular feature extraction and dimension reduction techniques used in the appearance-based approach. From classification point of view, generally, LDA-based algorithms perform better than that of PCA-based algorithms. However, there are two major problems when applying LDA in face recognition. The first one is the small sample size (S3) problem. It occurs when the image dimension is larger than the number of training samples. In turn, the within-class scatter matrix becomes singular. The second is the complicated and nonlinear image distributions when the images are captured under different poses and illuminations. In this thesis, discriminant analysis algorithms are designed and proposed, from the subspace approach and the regularization approach, to overcome these two limitations.

In the subspace approach, a new Subspace-LDA (SLDA) method is proposed to solve the S3 problem. Compared with the existing LDA-based face recognition methods for solving S3 problem, the proposed SLDA method is more efficient and gives better performance. By applying the kernel trick in the SLDA method, a Kernel Subspace-LDA (KSLDA) method is proposed to solve the nonlinear face image distribution problem. One of the crucial factors in the Kernel approach is the determination of the kernel parameters which highly affect the generalization performance of the kernel-based learning methods. To further improve the performance of our proposed KSLDA method, an automatic parameter estimation algorithm based on eigenvalue stability, namely the Eigenvalue Stability Bounded Margin Maximization (ESBMM), is developed. The proposed ESBMM algorithm is able to estimate multiple kernel parameters in RBF kernel function. Moreover, the ESBMM algorithm is generic and can be applied to other kernel-based LDA methods. To demonstrate its effectiveness, the proposed ESBMM algorithm is also applied to another kernel-based LDA method, namely the Kernel Direct-LDA (KDDA) method. Experimental results show

that after applying the ESBMM algorithm, the performance of the KSLDA method and the KDDA method are both improved by around three percents.

This thesis also explores the regularization approach to solve the above-mentioned two problems. by applying the kernel trick into one-parameter regularized discriminant analysis (RDA) method, a kernel-based one-parameter regularized Fisher discriminant (K1PRFD) method is proposed. A gradient decent-based algorithm is also developed to find the optimal regularization parameter.

In the final part, to utilize different advantages of different classifiers, a weighted combination scheme is proposed to integrate different kinds of face recognition algorithms. By integrating the concept of component-based approach and the weighted combination scheme, we proposed a component-based SLDA method to solve the one training sample problem for LDA-based algorithms in face recognition. A face image is divided into five sub-blocks. The SLDA method is then applied to each block. The final conclusion is drew based on the fusion results of the five SLDA classifiers. Two other practical applications also demonstrate the effectiveness of the proposed weighted combination scheme.

Each algorithm developed in this thesis has been extensive evaluated using public available databases such as FERET, YaleB and CMU PIE databases. Comparison between the proposed algorithms and the existing related algorithms are also performed and reported in this thesis.

In short, The major contributions of this thesis are summarized as follows:

- An efficient and effective SLDA method and its kernel version KSLDA method are proposed to solve the S3 problem and the nonlinear distribution problem;
- To improve the generalization performance of the kernel-based LDA methods, the Eigenvalue Stability Bounded Margin Maximization (ESBMM) algorithm is proposed to select kernel parameters for the kernel-based LDA methods;
- Integrating kernel trick into one-parameter regularized discriminant analysis algorithm, the kernel-based one-parameter regularized Fisher discriminant (K1PRFD) method is proposed to solve the S3 problem and the nonlinear distribution problem.

- A weighted combination scheme is proposed to integrate different kinds of face recognition algorithms. By integrating the concept of component-based approach and the weighted combination scheme, a component-based SLDA method is proposed to solve the one training sample problem.

Moreover, this thesis concludes that

- Both the subspace approach and the regularization approach are good and feasible approaches to solve the S3 problem and the nonlinear distribution problem in face recognition;
- In general, kernel-based methods gives better performance than that of linear methods, provided that the appropriate parameters are selected;
- Kernel parameter(s) selection is a crucial factor for kernel-based methods. The proposed ESBMM algorithm has demonstrated this issue.

# Table of Contents

Declaration	i
Abstract	ii
Acknowledgements	v
Contents	vi
List of Tables	xi
List of Figures	xiii
List of Symbols	xvi
List of Abbreviation	xvii
<b>1 Introduction</b>	<b>1</b>
1.1 Background . . . . .	1
1.2 Face Recognition . . . . .	4
1.2.1 Appearance-based Approach . . . . .	6
1.2.1.1 Principal Component Analysis . . . . .	6
1.2.1.2 Independent Component Analysis . . . . .	8
1.2.1.3 Linear Discriminant Analysis . . . . .	8
1.2.2 Model-based Approach . . . . .	9
1.2.2.1 Elastic Bunch Graph Matching . . . . .	10
1.2.3 Existing Problems . . . . .	11
1.3 Motivations . . . . .	12

1.4	Contributions of this Thesis . . . . .	13
1.5	Thesis Outline . . . . .	15
<b>2</b>	<b>Review on Existing LDA-based Algorithms and Public Available Databases</b>	<b>17</b>
2.1	Basic Theory of LDA Algorithm . . . . .	17
2.1.1	Some Definitions on LDA . . . . .	17
2.1.2	Traditional LDA Algorithm . . . . .	19
2.1.3	Kernel-LDA Algorithm . . . . .	21
2.1.3.1	Kernel Functions . . . . .	22
2.1.3.2	Kernel-based LDA Algorithms . . . . .	23
2.2	Limitations of Traditional LDA Algorithm . . . . .	25
2.3	Review on LDA-based Face Recognition Algorithms . . . . .	25
2.3.1	LDA-based Algorithms . . . . .	25
2.3.2	Kernel-based LDA Algorithms . . . . .	31
2.3.2.1	Review on Kernel-based LDA Algorithms . . . . .	31
2.3.2.2	Review on Kernel Parameters Determination . . . . .	33
2.4	Testing Databases Used in This Thesis . . . . .	34
2.4.1	ORL Database . . . . .	34
2.4.2	FERET Database . . . . .	35
2.4.3	YaleB Database . . . . .	35
2.4.4	CMU PIE Database . . . . .	36
<b>3</b>	<b>Subspace-LDA Algorithm for Face Recognition</b>	<b>38</b>
3.1	Introduction . . . . .	38
3.2	Proposed Algorithm . . . . .	38
3.2.1	Derivation . . . . .	39
3.2.2	Algorithm . . . . .	41
3.2.3	Analysis of the Algorithm . . . . .	42
3.2.3.1	Uncorrelated Constraint . . . . .	42
3.2.3.2	Complexity . . . . .	43
3.3	Experiment Results . . . . .	44

3.3.1	Results on FERET Database . . . . .	44
3.3.2	Results on YaleB Database . . . . .	45
3.3.3	Results on CMU Database . . . . .	47
3.4	Discussion . . . . .	47
<b>4</b>	<b>Kernel-Subspace LDA Method and ESBMM Algorithm</b>	<b>50</b>
4.1	Introduction . . . . .	50
4.2	Proposed Method . . . . .	51
4.2.1	Kernel Subspace-LDA Method . . . . .	51
4.2.1.1	Nonlinear Mapping from Input Space to Feature Space	52
4.2.1.2	Applying SLDA Method on Feature Space . . . . .	53
4.2.2	Eigenvalue Stability Bounded Margin Maximization Algorithm	55
4.2.2.1	Eigenvalue Stability analysis . . . . .	56
4.2.2.2	Procedure of Eigenvalue Stability Bounded Margin Maximization Algorithm . . . . .	61
4.2.3	Applying ESBMM Algorithm on Existing Kernel-based LDA Methods . . . . .	62
4.2.3.1	Applying ESBMM Algorithm on Kernel Subspace- LDA Algorithm . . . . .	62
4.2.3.2	Applying ESBMM Algorithm on Kernel Direct-LDA Method . . . . .	64
4.3	Experiment Results . . . . .	65
4.3.1	Performance Evaluation on the KSLDA Method . . . . .	65
4.3.1.1	Results on FERET Face Database . . . . .	65
4.3.1.2	Results on YaleB Face Database . . . . .	66
4.3.1.3	Results on CMU PIE Face Database . . . . .	67
4.3.2	Performance Evaluation on the ESBMM algorithm . . . . .	70
4.3.2.1	Kernel Parameters Determination . . . . .	70
4.3.2.2	Results on FERET Face Database . . . . .	72
4.3.2.3	Results on YaleB Face Database . . . . .	74
4.3.2.4	Results on CMU PIE Face Database . . . . .	77



4.3.3	Evaluation on Effectiveness and Complexity: Full Search vs ESBMM Algorithm . . . . .	77
4.4	Discussion . . . . .	79
<b>5</b>	<b>Regularized LDA for Face Recognition</b>	<b>81</b>
5.1	Introduction . . . . .	81
5.2	Proposed Kernel One-parameter Regularized Fisher Discriminant Method	82
5.2.1	Regularization of $Q_\varphi$ . . . . .	83
5.2.2	Formulating One Parameter Regularization . . . . .	85
5.2.3	The K1PRFD Algorithm . . . . .	86
5.2.4	Determine the Optimal Parameters $\Theta = (\theta, t)$ . . . . .	88
5.3	Experimental Results . . . . .	89
5.3.1	Results on FERET Face Database . . . . .	89
5.3.2	Results on YaleB Face Database . . . . .	89
5.3.3	Results on CMU PIE Face Database . . . . .	91
5.3.4	Computational Complexity . . . . .	97
5.4	Discussion . . . . .	97
<b>6</b>	<b>Combining Classifiers</b>	<b>99</b>
6.1	Introduction . . . . .	99
6.2	Review on Combination Theoretical Framework . . . . .	100
6.2.1	Combination Scheme: Product Rule . . . . .	100
6.2.2	Combination Scheme: Sum rule . . . . .	100
6.2.3	Combination Scheme: Max Rule . . . . .	101
6.2.4	Combination Scheme: Min Rule . . . . .	101
6.3	Proposed Combination Scheme . . . . .	101
6.3.1	Weighted Sum Rule . . . . .	101
6.3.2	Weight Determination . . . . .	102
6.3.3	Measurement Normalization . . . . .	102
6.4	Applications . . . . .	103
6.4.1	Component-based LDA Algorithm for Face Recognition . . . . .	104
6.4.1.1	Overview of the Proposed Method . . . . .	104

6.4.1.2	Construction of Five Component Bunches . . . . .	105
6.4.1.3	Combining Local Component Bunches for Classification	106
6.4.1.4	Experiment Results . . . . .	107
6.4.2	Combining Linear and Nonlinear Feature-based Algorithms for Face Recognition . . . . .	111
6.4.2.1	Methods for Combination . . . . .	111
6.4.2.2	Experiment Results . . . . .	112
6.4.3	Combining Local and Global Features for Face Recognition . .	116
6.4.3.1	Methods for Combination . . . . .	116
6.4.3.2	Experiment Results . . . . .	119
6.5	Discussion . . . . .	119
<b>7</b>	<b>Conclusions</b>	<b>121</b>
7.1	Summary . . . . .	121
7.2	Future Work . . . . .	123
	<b>Appendices</b>	<b>124</b>
<b>A</b>	<b>Derivation of Matrix <math>S_{wt}</math></b>	<b>124</b>
<b>B</b>	<b>Derivation of Matrix <math>Z</math></b>	<b>125</b>
<b>C</b>	<b>Derivation of Fisher Criterion</b>	<b>126</b>
<b>D</b>	<b>Expressions of <math>\nabla J = (\frac{\partial J}{\partial \theta}, \frac{\partial J}{\partial t})</math></b>	<b>129</b>
	<b>Curriculum Vitae</b>	<b>147</b>