

ระบอบตัวจำแนกหลายตัวที่มีนัยทั่วไปบนพื้นฐานหลักทฤษฎีสมมติฐานเฉพาะที่



นายวิทยากร อัครวิเศษ

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรดุษฎีบัณฑิต

สาขาวิชาวิศวกรรมไฟฟ้า ภาควิชาวิศวกรรมไฟฟ้า

คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

ปีการศึกษา 2548

ISBN 974-53-2038-2

ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

199200496

22 พ.ย. 2549

GENERALIZED MULTIPLE CLASSIFIER SYSTEMS WITH LOCAL  
DISCRIMINANT BASES

Mr. Widhyakorn Asdornwised

A Dissertation Submitted in Partial Fulfillment of the Requirements  
for the Degree of Doctor of Philosophy Program in Electrical Engineering

Department of Electrical Engineering

Faculty of Engineering

Chulalongkorn University

Academic Year 2005

ISBN 974-53-2038-2

Thesis Title                    GENERALIZED MULTIPLE CLASSIFIER SYSTEMS WITH LOCAL DISCRIMINANT BASES

By                                     Mr. Widhyakorn Asdomwised

Field of Study                  Electrical Engineering


Thesis Advisor                 Associate Professor Somchai Jitapunkul, Dr.Ing.

---


Accepted by the Faculty of Engineering, Chulalongkorn University in Partial Fulfillment of the Requirements for the Doctor's Degree

 ..... Dean of the Faculty of Engineering  
(Professor Direk Lavansiri, Ph.D.)

#### THESIS COMMITTEE

 ..... Chairman  
(Professor Prasit Prapinmongkolkarn, D.Eng.)

 ..... Thesis Advisor  
(Associate Professor Somchai Jitapunkul, Dr.Ing.)

 ..... Member  
(Suebskul Phiphobmongkol, Ph.D.)

 ..... Member  
(Assistant Professor Chedsada Chinrungrueng, Ph.D.)

 ..... Member  
(Assistant Professor Somchart Chokchaitam, D.Eng.)

วิทยากร อัครวิเศษ: ระบบตัวจำแนกหลายตัวที่มีนัยทั่วไปชนิดฐานหลักดิสคริมิแนนท์เฉพาะที่  
(GENERALIZED MULTIPLE CLASSIFIER SYSTEMS WITH LOCAL DISCRIMINANT BASES), อ. ที่ปรึกษา: รองศาสตราจารย์ ดร. สมชาย จิตะพันธ์กุล, 121 หน้า, ISBN 974-53-2038-2

ในหลายปีนี้ เทคนิคที่ประสบความสำเร็จและแพร่หลายสำหรับการรู้จำรูปแบบและจักรเรียนรู้ คือ ตัวจำแนกตระกูลเฉพาะ ที่รู้จักกันดีในนาม *ระบบตัวจำแนกแบบหลายตัว* ตัวจำแนกเหล่านี้ได้ถูกทดลองให้เห็นจริงในงานด้านต่างๆโดยสม่ำเสมอ ในบางกรณีสามารถปรับปรุงความแม่นยำในการทำนายให้มีผลดีขึ้นอย่างชัดเจน โดยเฉพาะเมื่อเปรียบเทียบกับระบบที่ให้ตัวจำแนกเพียงตัวเดียว วิทยานิพนธ์ฉบับนี้ขยายแนวคิดเดิมออกไปในหลายแง่ดังนี้ ปรับปรุงความถูกต้องในการรู้จำของระบบตัวจำแนกแบบหลายตัวโดยให้การแปลงเชิงเส้น การเข้ารหัสช่องสัญญาณที่มีนัยทั่วไปแบบคงทนสำหรับระบบตัวจำแนกแบบหลายตัว และการรวบรวมการทำนายอย่างเล็งเลศ

สาระสำคัญของวิทยานิพนธ์ฉบับนี้ได้แก่การสำรวจการนำมาใช้ของ *เฟรม* ซึ่งเป็นเซตเกินบริบูรณ์ของเวกเตอร์ เพื่อให้สำหรับการพรรณนาเชิงดิสคริมิแนนท์ที่มีประสิทธิภาพ ที่เรียกว่า *การขยายแบบเฟรมด้วยฐานหลักดิสคริมิแนนท์เฉพาะที่* (LDFE) ระเบียบวิธีที่นำเสนอเป็นแบบแผนที่ไม่สลับซับซ้อนและมีประสิทธิภาพเหมาะสำหรับปรับขยายการใช้อัลกอริทึมการสกัดลักษณะบ่งต่าง *ฐานหลักดิสคริมิแนนท์เฉพาะที่* (LDB) ให้อยู่ในเค้าโครงของ *การเข้ารหัสแบบหลายส่วนลักษณะ* (MDC) ในการเอาชนะต่อการจำแนกผิด ปริมาณความซ้ำซ้อนที่จัดสรรล่วงหน้าได้ถูกเพิ่มให้กับข้อมูลเดิมในระหว่างกระบวนการสกัดลักษณะบ่งต่าง การจัดสรรดิสคริมิแนนท์แบบไม่เท่าเทียมถูกสร้างขึ้นด้วยการแปรเปลี่ยนปริมาณความซ้ำซ้อนโดยให้มีนัยขึ้นกับความสำคัญของข้อมูล สาระที่สอง วิทยานิพนธ์นี้ได้พัฒนาส่วนขยายจำนวน 3 แบบจากวิธี ECOC ดั้งเดิมบนฐานของแบบแผนการเข้ารหัสต่อกันที่มีนัยทั่วไป การพัฒนาในส่วนนี้เป็นการพยายามในการปรับปรุงการจำแนกโดยการต่อกันของระบบตัวจำแนกแบบหลายตัวที่ต่างชนิดกันตั้งแต่สองตัวขึ้นไปเข้าด้วยกัน สาระสำคัญอีกส่วนหนึ่งของวิทยานิพนธ์นี้ได้แก่การหาผลลัพธ์ที่ดีที่สุดของการประกอบกันของการทำนาย โดยที่ได้ทำการสำรวจอัลกอริทึมใหม่สำหรับแบบแผนการรวบรวมแบบมีการถ่วงน้ำหนัก อัลกอริทึมนี้ให้การประมาณสันที่มีการปรับแต่งค่าพารามิเตอร์ทางสถิติ สาระสุดท้ายของวิทยานิพนธ์นี้ได้แก่การรู้จำหน้าด้วยฐานหลักดิสคริมิแนนท์เฉพาะที่ด้วยข่ายวงจรประสาท เนื่องจากความสามารถในการอินเทอร์โพลेटอย่างสูงของข่ายวงจรประสาท ทำให้สามารถสร้าง *กลุ่มของข่ายการแปลง* ซึ่งสามารถอธิบายได้ด้วย เค้าโครงการเพิ่มการเรียนรู้ (ในระดับเอาร์ทพุต) และการถ่วงน้ำหนักแบบจำลองเบย์เซียน

ภาควิชา ..... วิศวกรรมไฟฟ้า  
สาขาวิชา ..... วิศวกรรมไฟฟ้า  
ปีการศึกษา ..... 2548

ลายมือชื่อนิสิต .....  
ลายมือชื่ออาจารย์ที่ปรึกษา .....

##4471821021: MAJOR ELECTRICAL ENGINEERING

KEY WORD: MULTIPLE CLASSIFIER SYSTEMS / MULTIPLE DESCRIPTION CODING MODELS  
/ LOCAL DISCRIMINANT BASES / GENERALIZED CODE CONCATENATION

WIDHYAKORN ASDORNWISED: GENERALIZED MULTIPLE CLASSIFIER SYSTEMS WITH  
LOCAL DISCRIMINANT BASES, THESIS ADVISOR: ASSOCIATE PROFESSOR SOMCHAI  
JITAPUNKUL, Dr.Ing., 121 pp., ISBN 974-53-2038-2

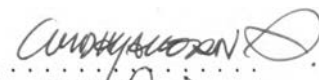
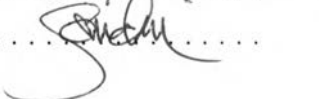
In recent years, the most successful and pervasive technique for pattern recognition and machine learning is a special family of classifiers known as *multiple classifier systems*. They have all demonstrated consistent – in some cases, remarkable improvements in predictive accuracy over single classifier systems. This dissertation extends the traditional concept in many aspects: improved recognition accuracy of multiple classifier systems through the use of linear transforms; robust generalized channel coding for multiple classifier systems; and optimal combining of predictions.

The first contribution of this dissertation is an exploration of the use of *frames*, which are overcomplete sets of vectors, to form efficient discriminant representations, called *Local Discriminant Frame Expansion* (LDFE). The scheme is a simple and efficient method suitable for extending the *Local Discriminant Bases* (LDB) feature extraction algorithm into *Multiple Description Coding* (MDC) framework. To combat misclassification, preassigned amounts of redundancy are added to the original data during the feature extraction process. Unequal discriminant assignment is implemented by varying the amount of redundancy with the importance of data. For the second contribution, this dissertation develops three extensions of the original ECOC method based on more generalized concatenated coding schemes. These are the attempts to improve classification through the concatenations of two or more heterogeneous multiple classifier systems. Other contribution is the optimization of combining of the predictions. In particular, a new weighted combining scheme is investigated. The algorithm utilizes a ridge estimator with statistically tuning parameter. Finally, a face recognition task is performed by using LDB with neural networks. Based on the key observation on the high interpolation power of neural networks, a collection of transform networks is constructed, in which it can be interpreted as both the frameworks of *incremental learning* (at the output level) and *Bayesian model averaging*.

Department . . . . . Electrical Engineering  
Field of study . . . . . Electrical Engineering  
Academic year . . . . . 2005

Student's signature . . . . .

Advisor's signature . . . . .

## Acknowledgements

I owe the most and greatest thanks to my advisor, Dr. Somchai Jitapunkul, for believing in me to pursue my scholastic goals. His support, critical reviews, discussions, patience, and, above all, the freedom were invaluable and caused this dissertation possible. In several occasions that I worked with him, he had shown how to solve many difficult engineering problems effortlessly. I was very fortunate to have an opportunity to work with him.

Particular thanks are due to Dr. Prabhas Chongsatiwattana of Department of Computer Engineering for many stimulating discussions and suggestions. He has always been ready for helps on various problems since the beginning through the end of the course of my dissertation. Thanks are also due to Dr. Chedsada Chinrungrueng. Many discussions with him contributed not only to the quality of the dissertation work, but also to the course of my life during the whole process of this research, especially when I had no way out.

I am also indebted to Professor Dr. Prasit Prapinmongkolkarn who provided guidance to me in several occasions in my academic life and professional. His kindness and passion in knowledge have consistently been shown to many students at the Department of Electrical Engineering, Chulalongkorn University. Thanks are also due to Professor Dr. Prasit Prapinmongkolkarn and Dr. Chedsada Chinrungrueng for serving as members on my dissertation committee.

I thank Dr. Suebskul Phiphobmongkol and Dr. Somchart Chokchaitam for serving as members on my dissertation committee. Their valuable discussions and insightful comments have improved the quality of this dissertation.

I would also like to acknowledge the scholarship from Ministry of Education, Thailand and the partial support from the Cooperation Project between Department of Electrical Engineering and Private Sector for Research and Development, Chulalongkorn University.

I sincere thank my parents for their support, encouragement, and love. I also wish to thank my daughter, Nathawan (Nathy), for creating wonderful distractions from my daily work. I was very fortunate to acquire the knowledge on several efficient learning and classification algorithms during the precious and very educational periods while I was baby sitting her. Last but not least, I would like to thank my wife, Usavadee. Without the love, support, sacrifice, and many life long discussions that I have consistently been received from her, none of this would be possible.

# Table of Contents

	Page
Abstract in Thai . . . . .	iv
Abstract in English . . . . .	v
Acknowledgements . . . . .	vi
Table of Contents . . . . .	vii
List of Tables . . . . .	x
List of Figures . . . . .	xi
 Chapter	
1 Introduction . . . . .	1
1.1 Subspace Methods of Pattern Recognition . . . . .	2
1.2 Motivations . . . . .	3
1.3 Overview of the Dissertation . . . . .	4
 2 Basic Background and Related Topics . . . . .	 6
2.1 Introduction . . . . .	6
2.2 Construction Approaches for Multiple Classifier Systems . . . . .	8
2.3 A Review of Some Multiple Classifier Systems . . . . .	8
2.3.1 Bagging . . . . .	8
2.3.2 Adaboost . . . . .	9
2.3.3 Error Correcting Output Codes . . . . .	9
2.3.4 Random Subspace Methods . . . . .	10
2.4 Recent Formal Definitions of Multiple Classifier Systems . . . . .	12
2.4.1 Classical Theory: Bias/Variance Dilemma . . . . .	12
2.4.2 Relevance to Optimal Encoding of Concepts . . . . .	13
2.4.2.1 Kolmogorov Entropy . . . . .	13
2.4.2.2 T-Repetition Coding Approach and Its Connection with Adaboost . . . . .	15
2.4.3 Relevance to Variance Reduction Techniques in Monte Carlo Methods 2.4.3.1 Importance Sampling . . . . .	16 17
2.4.3.2 Antithetic and Common Variates . . . . .	18
2.5 Rationale for a New Approach . . . . .	21
 3 Multiple Description Coding for Multiple Classifier Systems . . . . .	 24
3.1 Introduction . . . . .	24

Chapter	Page
3.2 A Library of Orthonormal Bases and Local Discriminant Basis . . . . .	25
3.2.1 Wavelet Bases, Wavelet Packet Bases and Best Basis Selection . . . . .	26
3.2.2 Local Discriminant Basis . . . . .	27
3.3 Multiple Description Coding Models . . . . .	28
3.3.1 Frame . . . . .	29
3.3.2 Biological Plausible Motivations . . . . .	29
3.3.3 Multiple Description Coding with Shift-variant Discrete Wavelet Transforms . . . . .	30
3.4 Local Discriminant Frame Expansions . . . . .	31
3.4.1 Redundant Discrete Wavelet Packet Transform . . . . .	32
3.4.2 Adapting LDB to Local Discriminant Frame Expansions . . . . .	32
3.5 Multiple Description Coding for Multiple Classifier Systems with Local Discriminant Frame Expansions . . . . .	37
3.6 Experimental Results . . . . .	37
3.7 Numerical Comparison . . . . .	44
3.8 Conclusions . . . . .	47
<b>4 Generalized Code Concatenation for Multiple Classifier Systems . . . . .</b>	<b>50</b>
4.1 Introduction . . . . .	50
4.2 Channel Coding . . . . .	52
4.2.1 Coding Theory . . . . .	52
4.2.2 Generalized Code Concatenation . . . . .	55
4.3 Applications to Multiple Classifier Systems . . . . .	59
4.3.1 Classical Concatenated Input-Output Code with Multiple Description Coding and Adaboost . . . . .	59
4.3.2 Classical Concatenated Input-Output Code with Multiple Description Coding and Block Codes . . . . .	59
4.3.3 Generalized Concatenated Input-Output Code . . . . .	60
4.4 Experimental Results . . . . .	62
4.4.1 Classical Concatenation Codes . . . . .	62
4.4.2 Generalized Concatenation Codes . . . . .	67
4.4.2.1 Multiple Description Ensembles of SVM-ECOC . . . . .	67
4.4.2.2 Bagged Ensembles of SVM-ECOC . . . . .	68
4.5 Conclusions . . . . .	69
<b>5 On Prediction Optimization Methods . . . . .</b>	<b>72</b>
5.1 Introduction . . . . .	72
5.2 Prediction Optimization Methods . . . . .	73
5.2.1 Majority Method . . . . .	73
5.2.2 Traditional Least Square Methods . . . . .	74
5.2.3 Antithetic Regression Method . . . . .	75
5.2.4 Ridge Regression . . . . .	76
5.2.4.1 Gradient Ridge Parameter Estimation . . . . .	77
5.2.4.2 Connection with Principal Component Method . . . . .	78
5.3 Discussion on Prediction Optimization Methods . . . . .	79
5.4 Diversity Measures . . . . .	80



Chapter	Page
5.5 Experimental Results . . . . .	82
5.6 Conclusions . . . . .	84
<b>6 Bayesian and Incremental Learning Frameworks . . . . .</b>	<b>88</b>
6.1 Introduction . . . . .	88
6.2 Local Discriminant Basis Neural Network Ensembles . . . . .	89
6.3 Multiresolution Committee of Networks : Bayesian Model Averaging Framework . . . . .	90
6.4 A Collection of Transform Networks . . . . .	94
6.5 Relation to Incremental Learning Framework . . . . .	96
6.6 Experimental Results . . . . .	99
6.7 Conclusions . . . . .	101
<b>7 Conclusions and Topics for Future Research . . . . .</b>	<b>104</b>
7.1 Conclusions . . . . .	104
7.2 Contributions of Dissertation . . . . .	104
7.3 Topics for Future Research . . . . .	105
<b>References . . . . .</b>	<b>107</b>
<b>Appendices . . . . .</b>	<b>116</b>
Appendix A . . . . .	117
Appendix B . . . . .	118
Appendix C . . . . .	120
<b>Vitae . . . . .</b>	<b>121</b>

## List of Tables

Table	Page
2.1 Coverage optimization methods. . . . .	7
2.2 Prediction optimization methods. . . . .	7
2.3 Error correcting output code: an exhaustive set of codewords. . . . .	11
2.4 Kolmogorov codebook. Maximum number of bits = $\log_2 N_\epsilon(\Xi)$ . . . . .	15
2.5 Relationship between common-antithetic variates and monotonicities of function $g$ and $h$ . . . . .	20
3.1 MSTAR images comprising training set. . . . .	40
3.2 MSTAR images comprising testing set. . . . .	40
3.3 The CGSM method : Recognition test of a three class problem for 80 x 80 images. . . . .	43
3.4 Multiple description pattern analysis using local discriminant frame expansions with 7 descriptions: Recognition test of a three class problem for 80 x 80 images. . . . .	43
3.5 Comparison of difference methods in overall percentage of images correctly recognized as a function of image size. . . . .	45
3.6 Comparison of training computational complexity. . . . .	48
3.7 Comparison of evaluation (testing) computational complexity. . . . .	48
3.8 Comparison of computational complexity for the MCS methods implemented in our experiment. . . . .	48
3.9 Comparison of accuracy (in overall percentage) and computational complexity for the MCS methods implemented in our experiment. . . . .	49
4.1 The performance of classical code concatenation method based on MDC and Adaboost for 3-class MSTAR data set. Comparison of difference methods is in overall percentage of images correctly recognized as a function of image size. . . . .	63
4.2 The performance of classical code concatenation method based on MDC and SVM-ECOC for 3-class MSTAR data set. Comparison of difference methods is in overall percentage of images correctly recognized as a function of image size. . . . .	66
4.3 The performance of generalized code concatenation method based on MDC and SVM-ECOC for Satimage data set. Comparison of difference methods is in overall percentage of patterns correctly recognized. . . . .	68
4.4 The performance of generalized code concatenation method based on Bagged ensembles of SVM-ECOC for Satimage data set. Comparison of difference methods is in overall percentage of patterns correctly recognized. . . . .	69
5.1 Comparison of different least square methods in overall percentage of images correctly recognized as a function of image sizes. . . . .	85
5.2 Comparison of different least square methods using various diversity measures. These are the best performances in overall percentage regarding to the optimal numbers of coefficients per description . . . . .	86
6.1 The performance of multiresolution committee of networks in overall percentage. . . . .	102

## List of Figures

Figure	Page	
2.1	The Adaboost algorithm. . . . .	11
2.2	Smooth function to be encoded. This could be the classification function of the observed data . . . . .	14
2.3	Best encoding of the smooth function with distortion $\epsilon$ . . . . .	14
3.1	Four methods of filtering and subsampling for one level decomposition discrete biorthogonal wavelet transform. a) Traditional wavelet transform. b) First alternate wavelet transform. c) Second alternate wavelet transform. d) Third alternate (left-shifting) wavelet transform. $c_0$ is a set of wavelet coefficients obtained from the previous level. $c_1$ and $d_1$ are the lowpass and highpass filtered signals subsampled by a factor of two. $z$ and $z^{-1}$ are the left and right circular shifting operators, respectively. . . . .	33
3.2	Results of four decomposition levels of a redundant wavelet packet transform and its simulated LDB functions. . . . .	33
3.3	Example of four transforms and their spatial discriminant levels. The gray area indicates the size of the most discriminant basis functions (or equivalently its power of discriminant). . . . .	35
3.4	Example of three spatially dispersed descriptions extracted by using local discriminant frame expansions of four transforms. . . . .	36
3.5	Multiple Description Pattern Analysis using Local Discriminant Frame Expansion.	39
3.6	Sample SAR images of military vehicles. (a) BMP2 APCs, (b) BTR70 APCs, and (c) T72 tanks. . . . .	39
3.7	One level decomposition of redundant versions of 2-D discrete wavelet transforms. a) to i) Transform 2 to Transform 10. $z_1$ and $z_2$ are the one unit horizontal and vertical circular left-shifting, respectively. . . . .	43
3.8	The performance of multiple description pattern analysis using local discriminant frame expansions at various image sizes. . . . .	45
3.9	The performance of multiple description pattern analysis using local discriminant frame expansions at different numbers of the most discriminant basis functions. a) 32 x 32 image size. b) 48 x 48 image size. c) 64 x 64 image size. d) 80 x 80 image size. . . . .	46
3.10	Notation for parameters used in complexity evaluation of MCS methods. . . . .	46
4.1	Code concatenation according to Forney. . . . .	57
4.2	Code concatenation according to Blokh and Zyablov. . . . .	57
4.3	Example of classical concatenation (systematic). . . . .	57
4.4	Example of codeword scheme for generalized code concatenation. . . . .	58
4.5	Example of generalized concatenated encoding. . . . .	58
4.6	Example of classical code concatenation based multiple classifier systems using MDC and Adaboost. . . . .	61
4.7	Example of classical code concatenation based multiple classifier systems using MDC and ECOC. . . . .	61
4.8	Generalized concatenation code scheme for multiple classifier systems. . . . .	63

Figure	Page	
4.9	Percent of recognition accuracy of the classical code concatenation method based on MDC and Adaboost for 3-class MSTAR data set. Each recognition accuracy is plotted as a function of number of weak classifiers and descriptions used in Adaboost at window size 32. . . . .	64
4.10	Percent of recognition accuracy of the classical code concatenation method based on MDC and Adaboost for 3-class MSTAR data set. Each recognition accuracy is plotted as a function of number of weak classifiers and descriptions used in Adaboost at window size 48. . . . .	64
4.11	Percent of recognition accuracy of the classical code concatenation method based on MDC and Adaboost for 3-class MSTAR data set. Each recognition accuracy is plotted as a function of number of weak classifiers and descriptions used in Adaboost at window size 64. . . . .	65
4.12	Percent of recognition accuracy of the classical code concatenation method based on MDC and Adaboost for 3-class MSTAR data set. Each recognition accuracy is plotted as a function of number of weak classifiers and descriptions used in Adaboost at window size 80. . . . .	65
4.13	Percent of recognition accuracy of the classical code concatenation method based on MDC and Adaboost for 3-class MSTAR data set. Each recognition accuracy is plotted as a function of image size and number of descriptions used. . . . .	66
4.14	Percent of recognition accuracy of the classical code concatenation method based on MDC and SVM-ECOC for 3-class MSTAR data set. Each recognition accuracy is plotted as a function of image size and number of descriptions used. . . . .	70
4.15	Percent of recognition accuracy of all of the code concatenation methods for Satimage data set. Each recognition accuracy is plotted as a function of image size and number of descriptions used. a) kernel width 32. b) kernel width 40. c) kernel width 48. d) kernel width 56. . . . .	71
5.1	Comparison of different combining methods with Q diversity measure at various window sizes with the overall percentage normalization to unity. a) 32 x32, b) 48x48, c) 64x64, and d) 80x80. <i>9ds_Maj</i> means that all 9 descriptions are used with majority combining. <i>Div_Maj</i> means that 7 descriptions are selected by using Q statistics. After selection, they are integrated by majority combining. <i>LS</i> , <i>Prin</i> , and <i>Ridge</i> are represented for traditional least square regressor, antithetic regressor with principal component approach, and ridge regressor with ridge parameter estimation, respectively. Note that each graph is plotted the high, low, upper-half standard deviation, and lower-half standard deviation of the recognition accuracy derived from various number of coefficients per descriptions. . . . .	83
5.2	Comparison of various methods for medium and high recognition accuracy with the overall percentage normalization to unity. Note that each graph is plotted the high, low, upper-half standard deviation, and lower-half standard deviation of the recognition accuracy derived from various number of coefficients per descriptions. a) Medium recognition accuracy. b) High recognition accuracy. Medium recognition accuracy means the prediction optimization method that has its mean and standard of deviation of the recognition accuracy in the middle range in term of the number of coefficients per description. This is similar to high recognition accuracy as well. . . . .	85

Figure	Page
6.1 A collection of neural networks trained by $N$ most discriminant subband images. . . . .	91
6.2 The Yale face database. . . . .	100
6.3 The first 10 most discriminant subbands. . . . .	100
6.4 Average recognition accuracy of the LDBNNE. The upper curves represent training results and the lower curves test results. . . . .	102