Automated analysis of human family relationship using facial features

Mohammad Ghahramani

2013

Mohammad Ghahramani. (2013). Automated analysis of human family relationship using facial features. Doctoral thesis, Nanyang Technological University, Singapore.

<https://hdl.handle.net/10356/55000>

<https://doi.org/10.32657/10356/55000>

Automated Analysis of Human Family Relationship Using Facial Features

Mohammad Ghahramani

School of Electrical & Electronic Engineering

A thesis submitted to the Nanyang Technological University
in fulfilment of the requirements for the degree of
Doctor of Philosophy

2013

Acknowledgements

I would like to express my sincerest gratitude to my supervisor, Associate Professor Eam Khwang Teoh, for providing me with crucial guidance and enthusiastic encouragement throughout the duration of this research work. His technical insights, literary and organizational expertise were crucial to the contents of this research work. I have learnt so much from his expertise and feel honoured to have the chance to work with him.

I would also like to thank my co-supervisor Dr. Wei-Yun Yau, currently the Program Manager at the Institute for Infocomm Research, I²R, and Adjunct Assistant Professor of School of Electrical & Electronic Engineering and Dr Yau really encouraged me in video surveillance systems. I greatly acknowledge him for his wise advice, ideas and his involvement in our research. He supported me with his guidance and collaborations in this research work. I also appreciate Dr. Karthik Nandakumar's assistance and the fact that he so generously shared his research work on Fusion algorithms.

Next, I would like to acknowledge the people whom have helped me through this journey in my personal life. I am forever indebted to my mother, Roya, my father Hossein Ghahramani, receive my deepest gratitude and love for their dedication and the many years of support during my undergraduate studies that provided the foundation for this work and my sisters, Yasmin and Ladan.

Last, but not least, I would like to thank my wife Fahimeh, who is my sincere and lovely companion during my life, for exhibiting extraordinary patience throughout my stay in Singapore, and providing me with love and encouragement to make this research work as painless as possible.

I thank all staff members and friends in Computer Vision Lab at the School of Electrical and Electronic Engineering, Nanyang Technological University for their kind support and help. Finally, I would like to acknowledge the School of Electrical and Electronic Engineering, Nanyang Technological University, for providing the award of a scholarship and the research facilities for my work in NTU.

Summary

There has been increasing interest on the use of video surveillance to analyze human biometrics, especially in places where the chance of occurring crime is high, such as in large cities. As an example, 2185 children are being reported missing daily in the United States. It has been rarely looked into belonging a person to the same family as a soft biometric modal. We term the ability to verify query images as family members “family (kinship) verification.” Several studies have been conducted on family photo albuming, but the problem of family verification has not yet been directly addressed and analyzed. The lack of comprehensive published family datasets is evidence that the matter has received very little attention to date. This thesis focuses on providing solutions for challenges of family verification as a recent application of computer vision. We propose frameworks to recognize members of the same family based on the facts and psychological studies on family members’ resemblance to each other. We also investigate the demanding case of missing/unknown member verification as a critical real-life application that will benefit from this research.

In the first major section of this thesis, we study true assumptions behind the resemblance of family members to one another, and the challenges of family verification. We develop traditional face recognition approaches to facilitate family verification by extraction of facial resemblance among family members. We investigate the very recent family verification problem, define scenarios and collect a dataset of family albums. Our aim is to recognize if the query sample belongs to the

same family. In the first stage, we investigate facial resemblance extraction among family members to perform family verification. Similarities of facial patches are utilized as multiple experts to recognize family members. Finally, our proposed inter-patch constraint-free algorithm of facial patches analysis determines the redundant patches among family members. On average, five facial patches (29% of the facial patches) are chosen to achieve the same accuracy obtained by using all patches for family verification. Consequently, the computational time is saved by 71% compared to utilizing all patches.

The second major section of our thesis is focused on family verification in real-time systems to obtain more discriminative features with the least dimension. Firstly, we minimize the quantization error of the Local Binary Pattern (LBP) operator through the incorporation of uniformly-sampled thresholds for LBP (UTLBP). Up to 3.4% improvement was achieved for family verification compared to various features in the literature, while the feature dimension is considerably smaller than the well-known Haar-like features (60%). Moreover, the proposed operator is more robust than the conventional operators, as tested on the large CAS-PEAL dataset against background, illumination, aging, and accessories changes. It achieves up to 8% performance improvement depending on the CAS-PEAL probe set, compared to the best operator in the literature. Then, we propose a novel redundant feature removal algorithm based on the fact that there is less similarity among the faces of family members than that of an individual's face samples. The proposed algorithms are then employed on the UTLBP feature operator to select the most informative set of thresholds to outperform the Haar-like features, with 20 times less feature dimension. The proposed approach converges 10 times faster than the state-of-the-art face recognition algorithms do.

The third contribution of the thesis is to utilize from each member's image segment resemblance to the entire family. Our proposed method is to incorporate the degree of resemblance of each individual member's patch to perform family verification.

Our analysis and the important consensuses of psychologists' studies reveal that the facial resemblance between family members differs from member to member and is facial patch specific. We propose to estimate the amount of resemblance for each patch based on the prior information of the member. As the score level fusion is preferred in the literature due to its wealth of information and low data dimension, we embed the available a priori information in the score fusion rule. The proposed method outperforms the state-of-the-art score fusion rules in all scenarios by 14%. Moreover, the resemblance estimation slightly outperforms human observation in the designated survey. The proposed method could also improve object family verification on the selected Caltech-256 database by 8%.

In summary, two automated frameworks for recognizing family members are proposed. The first framework utilizes UTLBP operator and the novel redundant feature removal provides fast and low computational algorithms suitable for real-time processing. The other framework of facial patch resemblance extraction among family members benefits from offline training with parallel computing capability. Finally, the enhancement of family verification can potentially solve critical applications of daily life to find missing children and unknown parents.

Table of contents

Acknowledgements	i
Summary	iii
List of Figures	xi
List of Tables	xv
List of Symbols	xvii
1 Introduction	1
1.1 Motivation.....	1
1.2 Related Work.....	3
1.3 Objectives	5
1.4 Original Contributions of the Thesis	7
1.5 Organization of the Thesis	9
2 A Literature Review	12
2.1 Introduction	12
2.2 Face Data Analysis	15
2.3 Holistic Appearance-based Face Analysis	16

2.4 Statistical Learning Approaches	18
2.5 Feature Extraction	18
2.6 Feature Types	20
2.6.1 Haar-like Wavelets	20
2.6.2 Gabor Wavelets.....	21
2.6.3 Scale Invariant Feature Transform (SIFT)	22
2.6.4 Local Binary Patterns (LBP)	27
2.7 Matching Algorithms.....	28
2.7.1 Adaptive Boosting.....	29
2.7.2 Mutual Boosting (info-boost).....	31
2.8 Mixture of Experts (MOE).....	33
2.9 Multi-biometric Systems.....	35
2.10 Fusion Techniques	37
2.10.1 Feature Level Fusion	38
2.10.2 Score Level Fusion.....	39
2.10.3 Continuous Score Fusion Techniques	41
2.11 Analysis of Family Members' Faces Data	44
2.12 Concluding Remarks.....	46
3 Our Proposed Framework for Family Verification and Resemblance Extraction	48
3.1 Introduction	48
3.2 Challenges of Family Verification.....	50
3.3 Family Dataset and Scenarios	53
3.4 Experimental Scenarios and Performance Evaluation	55

3.5 Family Verification Modeling	59
3.6 Family Verification Using Conventional Face Verification Approaches ..	61
3.6.1 K-Nearest Neighbors.....	62
3.6.2 Support Vector Machines	63
3.7 Benefits of Mixture of Experts for Family Verification	64
3.8 Data Usage for Experiments.....	65
3.9 Extraction of Family Facial Patch Resemblance	70
3.9.1 Facial Patch Template	72
3.9.2 Resemblance Extraction	72
3.9.3 Decision Making Rule.....	73
3.9.4 Selection of Important Patches.....	74
3.10 Experimental Results and Discussion.....	76
3.11 Concluding Remarks.....	80
4 A Novel Approach for Removing Redundant Feature Sets	82
4.1 Introduction	82
4.2 Our Proposed Approach for Family Verification	84
4.3 Enhanced Feature Operator for Family Members' Face Description.....	86
4.3.1 Local Ternary Patterns (LTP1).....	87
4.3.2 Local Triplet Patterns (LTP2)	87
4.3.3 Local Directional Patterns (LDP1).....	88
4.3.4 Local Derivative Patterns (LDP2).....	88
4.3.5 Our Proposed Uniformly-sampled Thresholds for LBP (UTLBP).....	89
4.3.6 UTLBP Operator Distinctiveness Analysis	90

4.3.7	UTLBP Sampling Step and Redundant Feature Sets	91
4.4	Conventional Redundant Feature Removal Methods.....	94
4.5	Valid Assumptions of Features' Information Analysis	95
4.6	Our Proposed Strategy for Redundant Feature Set Removal by (GA).....	97
4.6.1	Our Proposed Chromosome Structure and Cost Function for Genetic Algorithm.....	98
4.7	Our Novel Approach for Redundant Feature Set Removal	100
4.8	Experimental results and Disscutions	103
4.8.1	Performance Comparison Using Selected Features.....	104
4.8.2	Redundant Feature Set Removal Comparison	105
4.8.3	UTLBP Performance on Face Recognition	110
4.8.4	UTLBP Features Discrimination Power on Family Data.....	113
4.9	Concluding Remarks	115
5	A New Framework for Utilizing Similarity of Family Member's Image Segments	116
5.1	Introduction	116
5.2	Family Members' Facial Parts Similarity in Psychological Studies.....	119
5.3	Our Prposed Framework.....	121
5.4	Conventional Score Fusion Rules	123
5.5	Consolidation of Family Member's Facial Parts Resemblance	124
5.5.1	Estimation of Similarity	128
5.6	Experimental Results and Discussions.....	128
5.6.1	Split Faces into Patches	128
5.6.2	Feature Extraction	130

5.6.3	Score Fusion Rule.....	130
5.6.4	Utilizing Human Perception	134
5.7	Families of Objects.....	137
5.8	Concluding Remarks	140
6	Conclusion and Recommendations	143
6.1	Conclusion.....	143
6.2	Recommendations	146
6.2.1	Performance Improvement of Face Recognition.....	146
6.2.2	Recognition of Individual Family Members	148
6.2.3	General Object Recognition Enhancement Using Similarity of Members' Image Segments	149
	Author's Publications	151
	Bibliography	153

List of Figures

1.1	(a) A pair of twin sisters, Kian and Remea, born to a British couple who both had mixed-ethnics parents, adapted from [9]. (b) Michael Jackson's daughter, Paris adapted from [15] and (c) Mark Lester's daughter, Harriet adapted from [15].....	2
2.1	Face recognition steps adapted from [2].....	12
2.2	Taxonomy of face recognition algorithms based on feature operators, pose changes dependency for face analysis adapted from [4].....	19
2.3	Extended Haar templates adapted from [28].....	21
2.4	Feature description of key-points (a) Neighbourhood (b) Direction calculation (c) generation of feature vector descriptor adapted from [14]	24
2.5	Features of (a) the SIFT key-point descriptor of an object producing variable number of features (b) DSIFT results in equal number of features.....	26
2.6	Basic LBP Operator extracts the feature equal to $0111\ 01112 = 119$	28
2.7	Summary of feature selection methods adapted from [5].....	29
2.8	The AdaBoost pseudo code 1997 adapted from [8]	30
2.9	The InfoBoost algorithm adapted from [11].....	32

2.10 (a) Complex decision boundary that cannot be learned by linear or circular classifiers. (b) Multiple classifiers spanning the decision space adapted from [12]	34
2.11 (a) Distributed detection system with data fusion center. (b) Optimum data fusion structure adapted from [7]	40
3.1 Samples of face recognition and family verification for President Obama (first row) and his family members' face samples (second row)	52
3.2 Face database samples (a) face recognition CAS-PEAL database adapted from [13] (b) Family albums samples adapted from [16]	54
3.3 Samples of 4 selected family datasets with inevitable changes of occlusion, accessories, closed eyes, background, lighting, and slight pose and expression changes. Parent's samples are shown at the first row and their corresponding children at the 2 nd row.....	57
3.4 General block diagram of conventional face recognition methods	61
3.5 SVM Optimal Separating Hyper-plane adapted from [1]	63
3.6 Samples of cropped faces with respect to eyes positions of families 1 to 3 at each row, parent's samples are the first two samples and the rest of photos are their corresponding children	67
3.7 The overall framework of family verification using informative facial resemblance.....	71
3.8 (a) The proposed facial template and cropped facial patches (b) patch indexing	72
3.9 The Family verification performance for family datasets in defined tests.....	76
3.10 The importance histogram color map for family (a) 1, (b) 2, (c) 3, (d) 4, (e) 5 and test 1; (f) color map indexing; (g) feature points localization.....	80

4.1	A general Face Recognition block diagram [6]	84
4.2	The overall framework of redundant feature set removal steps	86
4.3	(a) Gabor Wavelets reliabilities (b) selected UTLBP thresholds reliabilities for family verification (Test 1) of a random family dataset. Blue bars are the feature sets with reliability of higher than 50%. Yellow bars correspond to feature sets with less than 50% reliability.	93
4.4	The Pseudo code of the Chair-Varshney rule implementation on GA selected feature sets' corresponding classifiers.....	101
4.5	The average family verification total error (Err) and False Negative Ratio (FNR) of selected features for Test 1, Test 2.a and Test 2.b using the conventional face recognition method shown in Figure 4.1	106
4.6	The total error for the set of b most informative sets for different features using the proposed two-stage redundant feature removal, PCA and GA..	109
4.7	A facial image divided into 5×5 regions (b) the weights set for the weighted χ^2 dissimilarity measure. Black squares indicate weight 0.0, dark gray 1.0, light gray 2.0 and white 4.0.	110
4.8	Some example images in CAS-PEAL-R1 Database adapted from [13]	112
5.1	(a) The average face of family members and their overall facial patch resemblance (b) each member's facial patch resemblance to be considered in the family verification algorithm (c) family members' facial patches split using the golden ratio mask template implemented in general.....	118
5.2	Aligned and Cropped family faces are split into regions (b) the golden ratio mask template indexing implemented in Chapter 3 (c) modified patches selected in this chapter.....	122

5.3 Pseudo code of the proposed method to utilize individual member's facial patch similarity.....	129
5.4 The average family verification total error (Err) and False Negative Ratio (FNR) of selected features for (a) Test 1, (b) Test 2.a and (c) Test 2.b on facial patches	131
5.5 The ROC curves for the proposed fusion approach and the fusion rule by Nandakumar et al. [10] for (a) Test 1, (b) Test 2.b (c) Test 2.b	133
5.6 (a) samples of Bush's family dataset samples for the designed survey (b) the survey questions and multiple choices answers samples.....	134
5.7 The ROC curves for the proposed fusion approach using human perception, the proposed resemblance estimation method and the fusion rule by Nandakumar et al. [10] for (a) Test 1, (b) Test 2.b (c) Test 2.b	136
5.8 The ROC curves of match scores fusion rule comparison for the proposed approach and Nandakumar et al. [10] using Caltech-256 (a) motorcycle and (b) car side-view datasets.....	141
6.1 Simplified 2-Dimensional manifolds comparison for (a) face recognition adapted from [2] (b) the i^{th} face manifold in a family separated from others by the red boundary.....	147

List of Tables

2.1	Feature description of key-points adapted from [3].....	25
3.1	Family specifications and their characteristics in our collected database	56
3.2	AdaBoost, SVM, KNN and LDA average performances (FNR, FPR, Err) for different tests.....	68
3.3	Majority voting on the outputs of AdaBoost classifiers for each individual family member for different tests in selected iterations	69
3.4	Majority voting performance on the outputs of multiple classifiers for different tests.....	69
3.5	Family facial patch resemblance shown by patch reliability measurement color map.....	77
3.6	Determination of b informative patches for family 3 classification, test 1, with the FNR error in both training (Tr) and testing (Te) sets and testing stage total error (Err).	79
3.7	Minimum number of patches needed to achieve almost the accuracy of all patches for family verification of families 1-5 and test 1.....	79
4.1	Image samples for feature extraction and comparison of the extracted UTLBP and the conventional LBP	92

4.2 Vector dimension of extracted features from images of size 20×24 (pixel depth = 256).....	105
4.3 Family verification average FNR and FPR comparison using the conventional face recognition algorithm shown in Figure 4.1 and the proposed individual feature set classifier fusion	107
4.4 The average computational time comparison of the proposed method and conventional face recognition in Figure 4.1 in different iterations.....	108
4.5 Recognition error rate for feature operators in each probe set of the CAS-PEAL database.....	113
4.6 Family verification error evaluation for selected feature operators.....	114
5.1 Performance comparison of the SVM classifier, the proposed method and the algorithm suggested in [10].....	132
5.2 Outputs of the proposed method and the proposed method and the algorithm suggested in [10] on selected samples of 3 family albums for the 3 different scenarios	137
5.3 Samples of each motorcycles and car side-view family members	139
6.1 Performance comparison of Face Recognition (FR) [6] and our recommendation to perform Family Verification prior to Face Recognition (FVFR).....	149

List of Symbols

W	Image width
H	Image height
μ	Gabor wavelet scale
v	Gabor wavelet orientation
σ	Standard deviation
$\ \quad \ $	the norm operator
z	Coordinates of a pixel
$k_{\mu,v}$	wave vector
$k_{\max}$	maximum frequency
f^v	spatial frequency between Gabor kernels
c	The center pixel
$s(x)$	The step function
g_c	Gray-scale value of the center pixel
g_p	Gray-scale value of the surrounding pixel
$F_g(s)$	Distribution function of the genuine data
$F_i(s)$	Distribution function of the imposter data
S	Score vector
M_j^i	j-th Member of family #i

X_I	The face sample of an individual
X_Q	The query face sample
$\bar{X}_{M_j^i}$	The set of all faces belong to M_j^i
$\bar{X}_I$	The set of all faces of individual I
C_1	Genuine family class
C_0	Non-family class
α	Strangeness value
ρ	Diversity
N^{ff}	false decisions when the data is a family member
N^{fn}	false decision when the data is a non-family member
N^{tf}	true decisions when the data is a family member
N^{tn}	true decisions when the data is a non-family member
$C(N,s)$	permutation of s out of N
Thr_R	reliability threshold
P^f	the probability of verifying family members correctly
P^{ff}	the probability of verifying family members incorrectly
$\cap_{ff}$	intersection of incorrectly classified family members of two
$f_{g,p}(S)$	joint density functions of the genuine data scores
$f_i(S)$	joint density functions of the imposter data scores

Chapter 1

Introduction

1.1 Motivation

There has been increasing interest in the use of video surveillance, especially in places with high security needs, such as locations in cities such as London, Sydney, Singapore, and so forth to reduce the crime rate, such as the rate of child abuse. There are several studies to improve automated human recognition using surveillance cameras. The popular approaches use face and gait recognition and multi-biometrics, or the fusion of several biometric modalities [2]. The most accessible and informative information of the human is the face in front of the camera. It has priority over other biometrics, since it is natural, nonintrusive, and easy to use. As shown in [17], face is given the first rank versus other biometric features in a Machine Readable Travel Documents (MRTD). The face carries various types of information about the identity and other properties of a person that not only belong to him, but also make him different from other groups of people. Human beings use face recognition everyday in social relationships, and the ability of human beings to recognize faces has been studied for many decades in the fields of psychology and computer vision [18].

Figure 1.1: (a) A pair of twin sisters, Kian and Reme, born to a British couple who both had mixed-ethnics parents, adapted from [9]. (b) Michael Jackson’s daughter, Paris adapted from [15] and (c) Mark Lester’s daughter, Harriet adapted from [15]

Nowadays, the facial feature extraction, and recognition can be performed in “real-time” [19] using artificial intelligence. However, the main challenges of automatic face recognition and analysis stem from a combination of multiple causes, such as changes in illumination, background, accessories, distance, or various properties of the human face such as skin color and the effects of aging.

Recently soft biometrics, a term that describes the human characteristics of age, gender, ethnicity and so forth, has been determined using facial data [20, 21]. Besides using the current soft biometric properties, one can explore if a query face belongs to a certain group of people having some similarities in appearance. Due to genetic similarities, there is facial resemblance among family members and the closest relations a person has will be to the members of his family. In this study, we limit our definition of family to immediate members such as parents and siblings. This thesis focuses on the development of the ability to recognize whether a group of people are related which can be addressed as “family verification.” Family verification not only encounters all variations in face recognition but also will include factors such as mixed ethnicities, high effects of aging, multiple age groups and unbalanced datasets that naturally exist in a family. To motivate our work, the example of a million-to-one chance black and white set of twins is provided in Figure 1.1(a). It is difficult for the human brain to spot the similarities between such rare twins and recognize them

as family members. Recognizing members of the same family became critical when Mark Lester claimed that Michael Jackson's daughter was his biological daughter, a claim that could potentially lead to a multi-million-pound fortune. The court refused to allow the DNA test for this case [22].

In this thesis, we focus on designing discriminate features and extract family facial resemblances to make the computational intelligence able to recognize family members from imposters. There are many current potential applications for family verification, including but not limited to the following:

- **Enhancement of consumer products:** Digital cameras are capable of focusing on faces instead of other irrelevant photo segments. In cases where there are a large number of faces in family photos, the camera could be adjusted to focus on family faces rather than on other faces. Moreover, automatic photo annotation, albuming software, social robots, and IPTV systems require distinguishing family members from imposters to interact accordingly.
- **Absence of training genuine faces:** On average 2,185 children are reported missing each day from shopping centers, amusement parks, and various other locations. It would be difficult and time wasting to ask families to bring photos of their missing children and have surveillance systems search for the child. However, it would indeed be possible to use family face samples to look for missing children in huge shopping centers [23] or to identify unknown parents.

1.2 Related work

To the best of our knowledge, there is no prior research focusing on family verification as defined in the field of computer vision, unless auto-tagging applications and general frameworks of face recognition applicable to family verification. In this

section, we briefly introduce general approaches for family verification and focus on the ones that are most pertinent to our work. These algorithms are divided into the categories of holistic and geometric face analysis, feature based learning approaches, and photo album annotation. We analyze face recognition methods and enhance each step based on the facts and true assumptions of family datasets to achieve higher accuracy, less computational load, and simpler structure.

Holistic and geometric face analysis: Geometric feature-based approaches were among the early solutions for face recognition [24]. In these types of approaches, several facial features, such as the eyes, nose, mouth, and chin are detected. Then, properties and relative distances among them are considered for recognition. Although they reduce the dimensions considerably, the facial feature detection algorithms have not been developed accurately, particularly in cases of significant pose and illumination variation [25]. On the other hand, holistic appearance-based methods [26] have the advantage of being stable against facial geometric features, but they are not accurate enough to consider non-linear variations of original complex face manifolds [27]. The only paper that considered family verification mostly employed geometric face analysis on the collected family database, and achieved about 72% accuracy in family verification, concluding that the eye gray-scale value is the most accurate feature for kinship analysis [12]. The database included 143 pairs of one child face samples and one sample of his/her corresponding parent.

Feature based learning approaches: These approaches learn from the training data by considering the variations of extracted features or facial appearance. Nowadays, many successful algorithms for face data processing are learning-based [2]. It is also highlighted that face recognition is highly nonlinear and nonconvex when compared to face detection [2]. Finding the intersection of family members' facial similarities makes the family class boundaries highly non-linear, due to higher variations in the family members faces, which share smaller and more nonlinear space in the face manifold. To overcome the nonlinearity, To overcome this nonlinearity, a local

appearance-based feature space, such as Haar [28], Gabor wavelet-based features [29] and local binary pattern (LBP) [30] can be employed on the image.

Photo annotation: With the widespread availability of digital cameras and social networks, huge numbers of photographs are being uploaded or stored in personal and public albums. The task is to annotate the photos based on the location, scene [31], event [31] and the individuals [32]. In an early work by Das and Loui, the Euclidian distance of faces was compared with each individual to recognize the person and his or her relatives [31]. Research works in this category improve the performance of annotation [31, 32], as well as social relationships such as friends or classmates, with the help of gender and Markov logics [33]. However, family identity and the ability to recognize and categorize various family members are missing in this field, despite the serious need. There has been little investigation of human perception in terms of these abilities by researchers in psychology [34, 35].

1.3 Objectives

The objectives of this thesis are the development of robust, discriminative feature operators and computationally efficient frameworks targeted towards automatically verifying immediate members of a given family. The design of such a model is guided by the overall philosophy that it should be applicable to datasets with varieties of photometric changes and unseen members with a minimal amount of performance drop. Mainly, this thesis attempts to answer following questions:

- Every family has common features and facial resemblances among its members. How to design feature operators to find relevant resemblance in a family and robust enough to against photometric changes?

-
- In what way can we remove excess selected features, as the similarity among family members is less than face recognition of an individual?
 - How to consolidate the available information of family resemblance optimally to perform family verification to improve the performance in case of small training data sample size and unseen member classification?

Our answers to these questions are embedded in robust, unified, and analytical family verification frameworks and are generally classified according to three major directions:

1. **Feature operator:** this direction is focused on extracting the available information from family members' faces. Although family members resemble each other due to genetic proximities, it is harder to describe the resemblance in a family than it is to describe resemblance among multiple samples of an individual. In this direction, feature descriptors are enhanced to represent the face more discriminately.
2. **Feature selection:** in this direction, we focus on removing redundant selected features due to the less prominent similarities of family members to each other than those of an individual to himself.
3. **Information fusion:** as the final stage of classification, we derive the match score fusion rule to combine the resemblance information in a family.

The remainder of this chapter explains the major contributions of this thesis in further detail.

1.4 Original Contributions of the Thesis

The original contributions of this thesis can be listed as follows:

- (1) Due to the popularity and usefulness of soft biometric modals to improve human identity recognition, we come up with a new soft biometric modal in computer vision, called family verification. A novel analytical and robust framework for the very new problem of family verification is presented, relying on direction 2 of Section 1.3 to select facial patches with the highest amount of similarity among family members. We define three different scenarios based on real-life applications to evaluate the proposed methods for family verification. Currently, there is no comprehensive family dataset available to specifically evaluate algorithms of family verification. As the first step, we gathered more than 45 family datasets, comprising on the average 120 images per family with more than 5400 face samples in total. The datasets were collected from available digital shared albums with inevitable photometric changes. Then, they were sorted, labeled with additional eye and mouth coordinates, gender and age group for further processing by other researchers. We studied the facts and true assumptions of family resemblance as the basis for the proposed framework. Family members are related based on genetic proximities and we need to detect facial resemblance among multiple members, rather than from two faces only, as done in the literature [36]. For this task, we propose facial patch resemblance extraction using our modified version of a Golden ratio mask [37] to split the face into meaningful patches. The patches with the highest reliability in family verification are then selected as the most similar face regions among family members. Moreover, to minimize the feature dimensions, patches with repeated similarity information are removed. Later, the query image individual patch classification results are consolidated [7] to make the final decision in terms of belonging to a specific

family. Consequently, the computational time decreases by 71% compared to using all facial patches.

- (2) We develop a novel family verification technique to recognize family members that answers the following questions: 1) Based on larger intra-class variability of family members' faces, are current feature operators discriminative enough to extract minute similarity features from faces? and 2) How can redundant selected features be removed in light of the fact that larger variations or less similarities between family members compared to face recognition leads to less discriminant features? In order to answer these questions, we propose a method to extract more information embedded in facial images by the more distinctive Uniformly-sampled Thresholds and to remove redundant feature sets so as to reduce the computational complexity. We propose a novel chromosome structure that enables the Genetic Algorithm (GA) to search among feature sets and to select the most informative set for the corresponding family album. Moreover, we propose to remove redundant feature sets using an inter-feature set analysis scheme. The overall method consequently outperforms the conventional face recognition methods with a significant reduction in feature dimensions after removing redundant feature sets for the collected family database. We also examined the proposed operator on the large CAS-PEAL dataset in order to compare its robustness against controlled conditions of accessory, aging, lighting, and background under standard tests provided in the database. The proposed UTLBP operator outperforms other similar features and enhanced operators in the literature, especially against with aging variations that occur in family datasets. The performance improvement is more significant in case of missing family members, for which, contrary to expectation, the performance drops by enlarging the LBP operator radius. Moreover, the proposed framework converges 10 times faster the state-of-the-art face recognition algorithms.

-
- (3) A novel match score fusion rule is presented that focuses on answering the following two questions: 1) what is the rank of the resemblance of the facial segment of each family member to the whole family? And 2) in what way can we consolidate the a priori information of each individual member's facial patch similarity in the final decision for family verification? We reviewed the psychological studies and valid facts of the resemblance of individual family member's facial parts. It is inferred that the resemblance differs from member to member and that it depends upon their corresponding facial patch. We propose a novel algorithm to firstly estimate the resemblance of every member's facial patch and then consolidate the estimated information into recognizing the family member. We also designed an online survey with which to test human perceptions of the amount of similarity among family members in selected datasets. The experimental results indicate that the proposed approach could estimate the similarity slightly more accurately than human perception does.
- (4) Finally, we describe the extension of previous parts of our contributions to recognizing objects by employing the similarity of image segments of individual object family member's to the whole family. We extend the term "family" to objects and benefit from improvements achieved in human family verification. We examine the proposed method in order to estimate the resemblance of every member's image segment and consolidate the information on selected families of objects from the Caltech-256 database. The proposed method could improve the recognition accuracy for families of objects, such as car's side-view and motorcycle recognition.

1.5 Organization of the Thesis

The remaining chapters are organized as follows:

Chapter 2 reviews the background material related to our understanding and the development of our work. The chapter begins with an introduction to face data complexity and face recognition methods, their applications, and challenges. Various face analysis techniques, feature operators, feature selection methods, matching algorithms, and fusion rules are discussed for face recognition. In the final section, we define the family verification problem, its challenges, and valid assumptions and describe related psychological studies.

Chapter 3 describes our proposed framework to deal with the family verification problem. The key solution to determine family facial resemblances in order to recognize family members is achieved by the proposed framework, regardless of classifier and feature operator. Moreover, redundant patches that carry less diverse information are removed in order to reduce the computational load.

Chapter 4 introduces the second part of our contributions, which includes an enhanced distinctive operator for face representation. It is followed by the redundant feature set removal to reduce the computational load and feature dimension. We also examine the proposed operator discrimination power and robustness in face recognition applications on the large CAS-PEAL dataset as well.

Chapter 5 presents our proposed method for employing from the resemblance of individual member's facial parts. We estimate every similarity of the present member's facial segment to the entire family and consolidate the available information in order to enhance family verification. Extensions of this part of our work is presented on a family of objects at the end of this chapter and examined on the Caltech-256 object database and is shown to lead to more accurate object classification.

Chapter 6 provides conclusion to the thesis and recommendations for future research work in this area.

Chapter 2

A Literature Review

2.1 Introduction

A face is a three-dimensional data which has lots of variations like, illumination, pose, expression, accessory occlusion and rotation. There are mainly two approaches for face representation, holistic, which the face is analyzed using statistical approaches and on the other hand feature based methods that extract local features from face images. Due to low performance of holistic methods researchers tend to apply feature based algorithms nowadays. A feature based face recognition system generally consists of multiple steps that are almost the same in various applications as shown in Figure 2.1.

Figure 2.1: Face recognition steps adapted from [2]

This chapter surveys the face representation methods, promising and vigorously researched approaches to computer-assisted face data analysis. According to representation models from observation data, we categorize two alternative techniques for face data analysis. In section 2.2, we study the complex, high dimensional face subspace and the possible variations occur in the face samples. The general framework of face recognition approaches including feature extraction, feature selection (dimension reduction) and matching algorithms are mentioned in summary. We explain the first category of face representation which primarily extracts a holistic appearance based representation of the whole face region in section 2.3. Compared to early face recognition methods [24] PCA-based approaches [26] lead to more accurate face recognition. These approaches perform recognition directly on naive face data representation and reduce the high dimensional raw face image into sub-spaces such as Eigenfaces [38]. Later, the most significant subspace from training images is selected for face recognition. The shortcoming of linear appearance-based approaches is their sensitivity against global changes of illumination or mis-alignment and more importantly learning non-linear boundaries of the non-linear face manifolds [2, 39].

In Section 2.4, the second category of face data analysis is explained, known as learning based methods. They play an increasingly prominent role in many face processing applications due to their robustness against photometric changes and high accuracy [40]. The widely promising potency of learning-based models for use in face data analysis stems from both prior information about face(s) and the provided variations in the training data across different individuals and hence, provides reliable recognition tools that are both robust and generic for various applications [2]. There are two main sub-algorithms in each statistical learning-based algorithm, feature level processing and matching (classification) algorithm as depicted in Figure 2.1. The feature operator makes the face class discriminative from other classes and classification engines learn the boundaries among extracted features. The feature operator transforms the naive image data to another feature plane which is usually of

higher dimension compared to the original image [28]. The features are designated to be robust against illumination changes [40], rotation [41], view-point and scale [42, 43]. The discrimination power of feature operators is evaluated by distance measurement between features of the genuine and imposter samples [30, 44, 45]. Famous feature extractors are studied in details in section 2.6 to get familiar with their advantages. Later, we benefit from these operators for the specific challenges in family verification and enhance them to increase their distinctiveness to extract the facial resemblance of members to one another.

As aforementioned, the extracted feature dimension is usually more than the original image. The learning algorithm generally includes an algorithm prior to learning to reduce the feature dimension. We study the principles of the Boosting [8, 11] feature selection algorithms in details. Learning the complex non-linear boundaries of face manifolds either requires strong classifiers to recognize the decision boundaries or we need to employ multiple linear classifiers to divide the problem and conquer the whole problem [46]. In section 2.7, we investigate various applications of the Mixture of Experts to employ them in the challenging application of the family verification in later chapters.

The performance of a single biometric system is limited to the enrolled individuals of that biometric trait and there are more similar individuals of a single biometric trait than individuals sharing similar multiple traits [47]. Multi-biometric systems, multi-sample, multi-instance, multi-algorithm, multi-modal and hybrid systems are analysed in terms of the available data of the genuine class and their shortcomings. The point is in a multi-biometric system, if combination of the provided information is not performed well, the overall performance may be even less than each of the employed single biometric systems [48]. Therefore, it is necessary to choose appropriate fusion rule at the proper stage to benefit from various sources of information as compensation of the extra computational load and memory space

caused by multiple evidences. Section 2.4 to 2.10 are the most important part of this literature so as to put our contributions in perspective.

2.2 Face Data Analysis

Face image analysis is a machine vision problem, in which a three-dimensional face data with possible photometric changes is to be classified. Aside from pre-processing stages (e.g. illumination normalization), face detection and face alignment are the necessary stages to provide a well-centered, normalized face image to restrict parameters affect the performance. Face components are the basic reference points to align faces and various methods have been developed to detect their exact location [49, 50]. Face processing encounters some major practical difficulties of the face data which are summarized as follows:

- (1) **Variability of Facial Appearance**, the appearance of a face is affected by several parameters like the facial pose (due to camera viewpoint), illumination and facial expression that produce large intra-class variations. Indeed camera specifications, various imaging parameters, aperture, exposure time and image sensor specifications increase these variations.
- (2) **Complex Nonlinear Manifolds**, the face manifold is highly complex especially in case of photometric, pose and expression changes. It is even stated that the changes of face identity is mostly smaller than variations of an individual's face [51]. This necessitates development of algorithms able to learn nonlinear boundaries for classification.
- (3) **High Dimensionality and few instances**, the next challenge is the generalization problem which arises from the fact that a face image of

80×95 resides in a 7600-dimensional feature space if more features are not extracted from the face. Depending on the database, usually the number of samples per individual is less than 10 which makes it difficult to learn the manifold.

Face discrimination depends on three main steps of feature extraction to represent the face texture, matching algorithms to differentiate between faces and finally the fusion rule to consolidate the results obtained from multiple sources of information if applicable. Note that, allse algorithms could be robust if they are applied on a properly normalized face in terms of geometry and photometric properties. In general, a face discrimination system compares the extracted feature vector of the query image against those of faces in the database. The matching result identifies which class the face belongs or the fusion technique combines the available information from multiple sources to make the final decision. Each of feature extraction, matching algorithms and fusion techniques could be a bottleneck in a face data analysis system if they are not optimized for the specific application and pattern specifications. We continue this chapter by reviewing approaches proposed in the literature and correspond each of three main stages of face data analysis.

2.3 Holistic Appearance-based Face Analysis

Approaches based on facial components relative geometric properties were among the first solutions to distinguish between faces used in early days [24]. The very low computational power of the computers did not allow high dimension and complex features to be extracted and processed for face analysis to tackle alterations in the face data. The rapid technology revolution in processing systems enabled researchers to analyze the rich information embedded in the appearance of the face.

A famous approach for appearance-based face analysis was the principal component analysis (PCA) that reduced the training face dimension by incorporating the eigenfaces [26]. PCA considers a number of eigenfaces from a set of genuine and imposter training face images. In this way, each face is constructed as the combination of eigenfaces to form a feature vector with less dimensionality than the original image with a wealth of information for facial analysis. In another approach, Linear Discriminant Analysis (LDA) [52], a discriminant feature subspace is built to distinguish “optimal” faces of different classes. These linear and holistic approaches are not accurate enough due to the following restrictions of their nature,

- (1) **Linear transformation,** These methods transform the high dimensional face data linearly to a low-dimensional subspace such as eigenfaces or fisherfaces. Hence, they are not able to handle the non-linear boundaries of face classes accurately. The nonlinear boundaries of face manifolds are lost in the projected subspace and it restricts the performance of these methods in the challenging problem of discriminating between faces.
- (2) **Distance measurement as the matching algorithm,** In the projected linear subspace the vector representation of the face is compared with the genuine and imposter class by employing Mahalanobis distance (of different orders). Since the distance between sample representations of an individual could be greater than that of the imposter, matching the represented face does not lead to distinguish between faces manifolds.
- (3) **Recognizing the unseen data,** These linear methods are enhanced by employing nonlinear kernel algorithms such as kernel PCA [53] to tackle with the nonconvex, nonlinear face manifold. The transformation from the original image is performed by the help of nonlinear kernels while keeping the

nonlinear face details. However, kernel methods may not perform well on the unseen data [2] that usually occurs in real life cases [15, 23].

The solution to overcome the nonlinearity of face manifolds and extracting useful information from the highly redundant raw image space is to use statistical learning algorithms on the extracted features [18]. The aim is to construct classification techniques able to learn difficult nonlinear classification and regression problems in the complex feature space. Despite of nonlinearity and nonconvexity reduction by employing normalization and feature extraction algorithms, the problem still needs to be solved by designing a matching algorithm able to achieve accurate performance. So a successful face processing system is in need of combining optimally designated stages for feature extraction, feature selection and matching algorithms.

2.4 Statistical Learning Approaches

The statistical learning approaches discriminate faces by learning from the pool of extracted features from training data. The features could be produced from the whole appearance or as linear combination of raw images. The overall performance of these approaches depends on the discrimination power of feature operators and the learning ability of the matching algorithm. In this section, we briefly introduce various approaches in the literature for feature extraction and matching algorithms.

2.5 Feature Extraction

The main purpose of feature extraction is to simplify face manifold into less nonlinear and nonconvex feature space that still allows keeping nonlinearity in the resulting space. Since facial features are to be compared with other faces

corresponding feature points, images must be firstly normalized geometrically and photometrically. The extracted features must be distinctive enough to distinguish face classes and robust against photometric changes, slight pose changes and aging effect.

The taxonomy of major face recognition algorithms in Figure 2.2 provides an overview of face recognition technology based on pose dependency, face representation, and features used for matching. The variety of feature operators for the vast applications of face analysis necessitates focusing on feature operators and matching algorithms related to the specific application of family verification. Our research is focused on pose-dependent, viewer-centered face images using feature-based methods for family verification. We study the most recent and useful features operators in the literature in the next section.

Figure 2.2: Taxonomy of face recognition algorithms based on feature operators, pose changes dependency for face analysis adapted from [4]

2.6 Feature Types

2.6.1 Haar-like Wavelets

Haar like features are extremely primitive local difference features used as digital image features in face and object recognition. The name is originated from Haar wavelets (1909 by Alfréd Haar), which has similarity to the step-wise function. Normally image intensities are stored as RGB or Gray-scale intensities at every pixel. However, it tends to be expensive in terms of computation to extract Haar features by employing each wavelet with different size and origin through the whole image. Viola and Jones suggested an alternative method that considered rectangular regions of the image and summed up the pixels in this region [19]. So for the given class model (e.g. face), the learning model could be constructed through learning the subsets that define a class.

A simple rectangular Haar like feature is defined as the difference of the sum of pixels of areas inside the rectangle, which can be at any position and scale within the original image. This modified feature set is called a 2-rectangle feature. Viola and Jones also defined 3 more types of rectangle features and 4-rectangle features in Viola Jones object detection framework. These values indicate certain characteristics of a particular area of the image. Each feature type can determine the existence (or not) of certain characteristics in the image, such as edges or changes in texture. The number of features derived from each template is quite large and depends on the image size. Recently, Pham et al. defined 25 types of extended Haar-like wavelets that generate huge amount of features for each image [28]. The appearance of extended Haar templates is depicted in Figure 2.3.

Figure 2.3: Extended Haar templates adapted from [28]

2.6.2 Gabor Wavelets

We can analyze the image in various resolutions while keeping spatial relations information by employing Gabor wavelets with different parameters. Gabor filter parameters can be adjusted in terms of orientation, amplitude and frequency [40]. A Gabor filter is in fact a Gaussian modulated sinusoid in the spatial domain and as a shifted Gaussian filter in the frequency domain in two dimensions. A Gabor wavelet is defined as,

$$\Psi_{\mu,\nu} = \frac{\|k_{\mu,\nu}\|^2}{\sigma^2} e^{-\frac{\|k_{\mu,\nu}\|^2 \|z\|^2}{2\sigma^2}} \left[e^{izk_{\mu,\nu}} - e^{-\sigma^2/2} \right] |k_{\mu,\nu} = k_\nu e^{i\phi\mu} \quad (2-1)$$

where z represents a 2-dimensional input point [40]. The parameters μ and ν are the orientation and scale of the Gabor kernel. “ $\| \quad \|$ ” specifies the norm operator, and σ refers to the standard deviation of the Gaussian window in the kernel. The wave vector $k_{\mu,\nu}$ is defined as,

$$k_{u,\nu} = k_\nu e^{i\phi_u} \quad (2-2)$$

where $k_v = k_{\max}/f^v$ and $\varphi_u = \pi u/8$, if 8 different orientations are chosen. $k_{\max}$ is the maximum frequency and “ f^v ” is the spatial frequency between kernels in the frequency domain. In the literature, 5 different scales and 8 orientations of Gabor wavelets are used, e.g. $v \in \{0, \dots, 4\}$ and $\mu \in \{0, \dots, 7\}$. Gabor wavelets parameters are adjusted as $\sigma = 2\pi$, $k_{\max} = \pi/2$, and $f = \sqrt{2}$. The Gabor wavelet extracts features of an image by convolving the image with the filter bank. Convolution of image I and a Gabor kernel $\psi_{\mu,v}$ is written as,

$$O_{\mu,v} = \Psi_{\mu,v} * I(z) \quad (2-3)$$

that is called Gabor feature. Since the output $O_{\mu,v}$ to each Gabor wavelet is a complex value with a real part $R\{O_{\mu,v}\}$ and an imaginary part $I\{O_{\mu,v}\}$, we use its magnitude to represent the Gabor features as

$$\|O_{\mu,v}\| = \sqrt{R(O_{\mu,v})^2 + I(O_{\mu,v})^2} \quad (2-4)$$

The complete set of Gabor wavelet is the vector of concatenated features extracted by convoluting each wavelet with the image $I(z)$ as $G(I) = \{O_{\mu,v}: \mu \in \{0, \dots, 7\}, v \in \{0, \dots, 4\}, z = (x, y)\}$ at every pixel point.

2.6.3 Scale Invariant Feature Transform (SIFT)

David Lowe proposed SIFT features in 2004, a successful feature extraction algorithm for image matching [42]. SIFT features are extracted from images to help reliable multi-scale matching between different views of the same object. The extracted features are robust to scale and orientation changes and could accurately discriminate images. This method mainly includes four steps:

- (1) **Multi-scale feature extraction**, The first step is to convolve the image $I(x, y)$ in different scales by the Gaussian convolution kernel $G(x, y, \sigma)$, the pixel's coordinate of the input image is represented by (x, y) ; σ is the scale factor resulting the image description $L(x, y, \sigma)$ as,

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y) \quad (2-5)$$

where the Gaussian convolution kernel $G(x, y, \sigma)$ is

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2} \quad (2-6)$$

Using the resulting different scale images the interest-points (key-points) are detected in the way that the Difference of Gaussian (DoG) image between scale ' $k\sigma$ ' and ' σ ' is evaluated as,

$$\begin{aligned} DoG(x, y, \sigma) &= (G(x, y, k\sigma) - G(x, y, \sigma)) * I(x, y) \\ &= L(x, y, k\sigma) - L(x, y, \sigma) \end{aligned} \quad (2-7)$$

Image features, such as contour, corner and edges can be detected in different scales, and the possible image key-points are chosen by searching for the features extreme value. The key-points are extracted as features if the value of the pixel is less or higher than the existing 28 neighbors of other 2 octave scales.

- **Interest-point (key-point) location**, numerous key-points are detected in the previous step and those are sensitive to noise or have no edge or corner effect in this process must be removed. According to Taylor quadratic expansion, $DoG(x, y, \sigma)$ can eliminate the extreme points which have less contrast.
- **Key-point feature direction calculation**, after positioning and local extreme points removal the next step is to find the key-point's direction, which

ensures the feature's rotation invariance. The direction is calculated by the image information of key-point's neighborhood.

- **Key-point feature descriptor generation,** The position, scale and direction of key-points can only ensure the 2D geometric invariance, but it still needs to be robust against illumination and view transformation changes that is done by the SIFT descriptor. In the 16×16 pixel field, ' M ', in the neighborhood of key-point, split ' M ' into 16 region of size 4×4 . Then according to previous step the direction and amplitude in every region is calculated, in predefined 8 directions that is equivalent to $(2 * \pi / 8)$ as shown in Figure 2.4 (b).

Figure 2.4: Feature description of key-points (a) Neighbourhood (b) Direction calculation (c) generation of feature vector descriptor adapted from [14]

Finally the feature description is obtained by connecting the direction descriptions of all subfields. Therefore total number of direction descriptions is expressed by a $1 \times (16 * 8) = 1 \times 128$ vector as shown in Figure 2.4 (c).

The feature descriptor is then normalized to make the features invariant to illumination. In this way the SIFT descriptor is robust against scale, rotation, and affine variations and is highly robust against occlusion as claimed by D. Lowe [42].

He proposed an image matching equivalent to the feature matching to find the Euclidian distance between key-points from two pictures. The two key-points with minimum distance are chosen as the candidate of key-point matching pair. Finally, the ratio of the minimum distance and the second minimum distance is calculated and compared with the preset threshold. If the ratio is greater than the threshold then the query is selected as the potential candidate for matching. Finally, if the number of nearest key-point is greater than the 2nd preset threshold, the two images are matched [42].

Aside from SIFT feature advantages the shortcoming of this operator is the minimum image size due to octave scales of the image. In case of small image size, number of SIFT features drops significantly and the features are not unique for human identification. A detailed report on number of SIFT features of different images scales on selected face databases shows the effect of various distance measurements such as Euclidean, City-block and Cosine [3]. M. Aly reported the trade-off between number of SIFT features and face recognition performance on selected face databases as tabulated in Table 2.1. When the average number of SIFT features is between 30 to 80, there is an acceptable trade-off between feature extraction complexity and face recognition performance. However, variety of number of features for image sizes restricts classification algorithms for SIFT feature comparison and matching.

Table 2.1: Feature description of key-points adapted from [3]

	AT & T			Yale		
Resolution	#SIFT	Cosine	Angle	#SIFT	Cosine	Angle
100%	70	93.8	96.3	230	85.9	91.7
75%	53	92.6	95.8	155	87.2	91.4
50%	30	94.7	94.9	87	87.9	92.5
25%	10	88.4	88.4	33	79.2	82.8

Wang et al. proposed to extract descriptors around uniform image sampling points instead of only original SIFT key-points [54]. The resulting feature vector would be of equal size for normalized cropped images which is called Dense SIFT (DSIFT). The descriptors at the uniform sampling points are then concatenated to form the final vector. Other than parameters involved in the SIFT operator [42], step size of the sampling points in both vertical and horizontal directions need to be specified [54]. The position of key-points of an object's features extracted by the SIFT operator and the modified DSIFT operator uniform image sampling points are illustrated in Figure 2.5 (a) and Figure 2.5 (b) respectively

Figure 2.5: Features of (a) the SIFT key-point descriptor of an object producing variable number of features (b) DSIFT results in equal number of features

They achieved higher accuracy by fusing the DSIFT feature and the global shape context of the face image in the real time gender recognition application [54]. The Adaptive Boosting (AdaBoost) [55] was employed as the matching algorithm.

2.6.4 Local Binary Patterns (LBP)

Among feature operators in the literature, LBP feature vector dimension is the lowest and is robust to illumination changes. It has been successfully applied to face detection and recognition [56]. It has proven to be highly discriminative and its key advantages, specifically being invariant to monotonic gray level changes and computationally inexpensive, make it suitable for challenging image analysis and classification problems. The original LBP operator assigns a binary code to every pixel of an image by thresholding the surrounding neighbors with the center pixel value and the resulting binary code is calculated. The histogram of all pixels binary codes results in 256 bins if a 3×3 window is considered. Ojala et al. extended the LBP operator to different sets of sampling points, P , on the circle centered at the center pixel with radius R [41].

The LBP feature of pixel ' c ' is evaluated by employing the function ' s ' on the surrounding pixels around the center pixel. As in the literature, we define the signed difference function $s(x)$ as in equation (2-8).

$$s(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases} \quad (2-8)$$

Therefore the LBP feature is calculated for the center pixel with the value ' g_c ' and the value of surrounding pixels of ' g_p ' as,

$$LBP_{P,R}(c) = \sum_{p=0}^{P-1} s(g_p - g_c) 2^p \quad (2-9)$$

An illustration of the basic LBP operator (R,P) of (1,8), radius of 1 and neighborhoods of 8, is shown in Figure 2.6.

Figure 2.6: Basic LBP Operator extracts the feature equal to $0111\ 0111_2 = 119$

There are several improvements in the literature for the LBP operator trying to embed more facial information to the operator that will be discussed in details in Chapter 4.

2.7 Matching Algorithms

Feature extraction methods provide details of the image in a high dimensional subspace. Most of matching algorithms include feature selection methods to select dominant features. Feature selection or subset selection is a process commonly used in machine learning, where a subset of discriminate features is selected from a pool of extracted features. The best subset contains the least feature dimension that contributes the most to accuracy without any excess data. This is an important stage after which a classification algorithm is implemented to find the class boundaries. There are two categories of feature selection, Forward Selection, FS, starting with no variables and add them one by one or group by group, at each iteration considering a criterion such as classification error. The second category the Backward Selection (BS), starts with all features and removes them in the same manner [5]. A complete taxonomy of algorithms is plotted in Figure 2.7 [5].

Figure 2.7: Summary of feature selection methods adapted from [5]

Bang-Jensen et al. discovered that greedy algorithms may not reach the optimal set or even diverge [57]. As a result, we consider matching algorithms and feature selection methods that have been proved to be stable and could learn complex boundaries of classes accurately.

2.7.1 Adaptive Boosting

Adaptive Boosting is a machine learning algorithm that can be used in conjunction with many other strong classifiers to achieve high accuracy. The idea of boosting was first proposed by Schapire [58]. He proved that any weak learning algorithm may be boosted to a strong one based on a Probably Approximately Correct (PAC) model, assuming that there exist weak learning algorithms which can do slightly better than random guessing [59]. The most famous boosting algorithm, Adaptive Boosting or AdaBoost, was introduced by Freund and Schapire in 1997 [8]. It is of much interest since the training error converges exponentially and it continues to learn even after the training error is zero as claimed by the authors and addressed in the literature.

On the other hand, it tries to generalize the classification to samples since it does not maximize the margin to prevent over-fitting. The pseudo code for AdaBoost is illustrated in Figure 2.8. It preserves a probability distribution $P_n(\mathbf{x})$ over the training samples of size 'm'. In each iteration 'n', it picks a training set by sampling with replacement according to the selected probability distribution $P_n(\mathbf{x})$.

Figure 2.8: The AdaBoost pseudo code 1997 adapted from [8]

The learning algorithm is then employed to generate a weak learner $h(n)$ by finding threshold for discriminant features. The error rate $e(n)$ of this classifier on the training samples is computed and used to increase the weight for error samples and update the $P_n(\mathbf{x})$. The point is AdaBoost is continuously adding the new classifier (feature sets) to the previous set of weak classifiers (selected feature subset) if and only if the error is less than 0.5. Obviously, it is not considering the set of features to have the best performance with the least dimension.

2.7.2 Mutual Boosting (info-boost)

In this algorithm during the AdaBoost learning process, mutual information between the candidate weak classifiers and the new selected weak classifier is examined in order to reduce feature dimension [11]. Therefore, the non-informative classifiers carrying information already captured by the previous feature/classifiers will be excluded. The process is forward selection and the extra computation required is very low in comparison to the original AdaBoost.

As a basic concept in information theory, entropy of X or $H(X)$ is generally used to measure the uncertainty of a random variable X . The mutual information (MI), $I(X;Y)$, of two discrete random variables X, Y is can be defined as,

$$I(Y;X) = H(X) + H(Y) - H(X,Y) \quad (2-10)$$

where,

$$H(X) = - \sum_x p(X=x) \lg(P(X=x)) \quad (2-11)$$

$x \in \{1, -1\}$, in AdaBoost algorithm. MI can be extended using Bayes rule as,

$$I(Y;X) = H(X) + H(X|Y) = H(Y) - H(Y|X) \quad (2-12)$$

If a new classifier is added to the previous selected classifiers, the MI of the new classifier and the strongest classifier is calculated. If it is more than a predefined Threshold of Mutual Information (TMI), then the new classifier is added. In case of multiple strongest classifiers, the max value of their MI is compared with TMI. Figure 2.9 illustrates the pseudo code for InfoBoost algorithm.

```

Given  $M$  training samples  $(x_i, y_i), i = 1, 2, \dots, M, x_i \in R^N, y_i \in \{-1, +1\}$ 
Initialization: weights  $w_1(i) = 1/M$ 
For  $t = 1, \dots, T$ 
  1) Train weak learners using distribution  $w_t$ 
  2) Given each candidate weak classifier  $h_j$ , calculate classification error

$$\varepsilon_j = \sum_i w_t(i) |h_j(x_i) - y_i|$$

  LoopCounter = 0
  Do
    Choose  $h_u$  with lowest error  $\varepsilon_u$  from the candidate classifiers
    Calculate the max MI  $R(h_u)$  according to Eq. (6)
    If  $R(h_u) < TMI$ 
      The classifier found,  $h_t = h_u, \varepsilon_t = \varepsilon_u$ 
      go to 3)
    Else
      Remove  $h_u$  from the candidate list
    End If
    LoopCounter = LoopCounter + 1
  While LoopCounter < MAX_LOOPS
  3) Calculate  $\alpha_t = \frac{1}{2} \ln \left( \frac{1 - \varepsilon_t}{\varepsilon_t} \right)$ 
  4) Update weights:  $w_{t+1}(i) = \frac{w_t(i) \exp(-\alpha_t y_i h_t(x_i))}{Z_t}$ 

Final strong classifier:  $H(x) = \text{sign} \left( \sum_{t=1}^T \alpha_t h_t(x) \right)$ 

```

Figure 2.9: The InfoBoost algorithm adapted from [11]

Estimation of MI needs the value of marginal distribution $p(X)$, $p(Y)$ and the joint probability distribution $p(Y, X)$ which are calculated by counting the number of possible cases and divided by the total number of training samples. The experimental results show that InfoBoost achieves lower training error rate with fewer classifiers. The performance of the InfoBoost reaches to a constant level of improvement from a number of weak classifiers onwards in selected experiments conducted in [11]. Threshold adjustment is one of the drawbacks of the mutual boosting.

2.8 Mixture of Experts (MOE)

There are many scenarios in which we consider decisions of multiple experts in our daily lives such as asking additional opinions before making a decision, asking different doctors' opinions, reading user reviews before buying a product, or requesting references before hiring someone. The main purpose of Mixture Of Experts, (MOE), is to improve our confidence for the correct decision, by the major vote, weighing various opinions, and combining them through some thought process (algorithm) to make the final decision. Similarly, ensemble systems improve the performance of the whole system based on a representative training data, for which the correct decisions are *a priori* known [46]. Based on the literature of MOE applications, there are several theoretical and practical reasons why we may prefer an ensemble system or when MOE improves the system performance [46].

- Statistical reasons

The training data must be representative of the overall data so that it covers all variations such as photometric changes and occlusion for face images. However such data is not accessible in most of applications. In other words, good performance on training data does not predict good generalization performance. It is more sensible for cases of data not seen during training as in unknown family members through the history [12] and missing children [23]. MOE is employed to benefit from different generalization performances of a set of classifiers with similar training performances.

- Too little data

In some applications enough representative set of training samples may not exist to train the classifier. In the absence of adequate training data, re-sampling

techniques can be used for drawing overlapping random subsets of the available data, each of which can be used to train a different classifier, creating the ensemble [60, 61]. Family photo datasets is one of the applications of inadequate datasets [12] which need to benefit from MOE.

- Divide and conquer

In some cases, the available data is representative enough for training but it is difficult for a given classifier such as linear classifiers to solve. The decision boundary that separates data from other classes may be too complex, or lie outside the space of functions that can be implemented by the chosen classifier model. i.e. considering linear classifiers for an elliptical boundary depicted in Figure 2.10.

Figure 2.10: (a) Complex decision boundary that cannot be learned by linear or circular classifiers. (b) Multiple classifiers spanning the decision space adapted from [12]

- Data type fusion

If we have several sets of data obtained from various sources, scenes, lightings and poses where the nature of features are different, a single classifier encounters difficulties to learn the information contained in all of the data. In the application of family verification each family member is a different sub-class of the main class or family that represents the final class.

2.9 Multi-biometric Systems

Conventional authentication methods based on passwords or (electronic) tokens were not able to distinguish the fake user. The current biometric systems are superior to the traditional systems in terms of accuracy, speed and there is no need of manpower to analyze the trait. Biometric systems use the appearance and characteristics of the person to identify the human being. A verity of applications such as borders checking, financial institutions, health care and various access systems are involved with biometric systems nowadays. However, the efficiency of a single biometric identifier depends on its robustness to various sources of attacks and fake genuine samples. Since the number of unique individuals with expected variations in the biometric trait is inevitable such as identical twins for face recognition the accuracy of a single biometric system would have an upper bound. These natural restrictions can be humbled using the available information of multi-biometric systems [47]. Using multi biometric systems was firstly addressed in [62] known as layered biometrics. The term multi-biometrics indicates the consolidation of different types of evidence. These systems are divided to several main categories. Those implemented on single modal biometric systems are explained in summary as below,

- Multi-sensor Systems

In general, multi samples are acquired by multi-sensors to obtain diverse pieces of evidence from a single biometric trait. Using multiple CCD cameras to obtain the three-dimensional face image to perform accurate face recognition [63] or with two different sensors of infrared thermal and visible light sensor [64]. Applications encounter insufficient number of training images with the possibility of obtaining multiple evidence of the biometric trait could benefit from the idea of multi-sensors.

- Multi-algorithm Systems

The same data acquired from the single biometric trait is analyzed using multiple algorithms to reduce the cost and hassle of multi-sensor systems. It needs to investigate multiple feature operators and/or matching algorithms that leads to higher computational load. The point is, if the information and classification results obtained from multiple algorithms are not diverse enough it only increases computational load and the performance may deteriorate compared to each algorithm alone. The multi-algorithm systems are easy to study since the required data can be generated by employing desired algorithms on the naive biometric trait.

- Multi-instance Systems

These systems benefit from multiple instances of the same biometric modal. Multiple fingers' fingerprints or the left and right irises are the possible applications of this method. They are efficient in terms of the cost due to single sensors and computational load as they do not require algorithms for data

processing. However, in some cases, multi sensors may be used for the convenience of the user or saving the time for capturing the data.

- Multi-sample Systems

In these systems a sensor(s) is/are used to obtain multiple samples of the single biometric modal to capture representative possible variations of a single trait from an individual. Capturing face images from multiple video frames or different poses using mirrors are the applications of these systems. The final data can be a concatenated form of the captured images or separate instances of the image processed individually.

There are two other categories of multi-biometric systems known as multi-modal and hybrid systems that need multiple sensor types to acquire the traits and their applications are restricted due to unavailability of multiple traits in most scenarios. Currently, there are large but few available databases of multi-sensor, hybrid and multi modal datasets [65] that restricts expanding algorithms in these categories to other human identification applications and the provided data are acquired from limited biometric traits.

2.10 Fusion Techniques

In a multi biometric system, it is crucial to combine the available information appropriately in order to prevent extra computational load or deteriorate the overall performance [66]. The method of merging the information or evidence provided by multiple biometric systems is known as information fusion. The goal is to find the best set of experts such as features, classifiers or instances and develop an optimal rule to consolidate the decisions delivered by each expert. The amount of information

from the query sample that enters the system keeps getting more and more condensed as we move from the acquisition sensors to the final matching algorithm that identifies the final class. In a multi-biometric system the information fusion is possible in different stages of a pattern recognition system. Fusion techniques are mainly divided into (i) feature level fusion and (ii) score level fusion [48]. The significant information compression occurring at the matching algorithm output causes considerable change of strategy to consolidate the available information.

2.10.1 Feature Level Fusion

The first possible stage of a pattern recognition system for information fusion is at the sensor level. The raw data is concatenated to form a vector of the acquired images for further processing. It is also called fusion at the pixel (image) level. It is suitable for multi-sample systems to mosaic the acquired data from each sensor. The next possible stage of fusion techniques is at the feature level. Combining extracted features by employing multiple operators on acquired biometric data of the same individual provides a more diverse representation of the biometric trait(s). However, unknown relation between the feature spaces of multiple biometric systems, irreconcilable feature vectors size and type of multiple trait/operators, computational complexity, curse of dimensionality and unavailability of feature level data of the commercial biometric devices are the challenges of feature level fusion. One needs to design appropriate feature normalization method to combine feature vectors of different distribution scale and nature as feature level fusion may not essentially improve the performance compared to utilizing each feature set alone [48].

2.10.2 Score Level Fusion

The matching module downsizes the available information significantly to a discrete or continuous vector/scalar. Fusion is utilized at the decision level or its thresholded value as many commercial biometric matchers provide access only to the final matching score. General logical/algebraic operators such as “AND”, “OR”, “Mean”, “Median”, “MAX/MIN” and “Weighted Majority Voting” rules could be employed on the output produced by the matcher(s) [46]. One of the shortcomings of logical rules is their extreme performance operating points. In a Receiver Operating Characteristic curve (ROC) they either achieve very low False Acceptance Rate (FAR) and high False Rejection Rate (FRR) more than each individual matcher or the vice versa [48]. Further analysis of fusion techniques require considering the fact that the matcher output is discrete or continuous that leads to distinct strategies to perform optimal fusion.

- Sub-Optimal Discrete Fusion rules

The matchers’ output is generally compared with the predefined threshold to obtain the final decision [67]. Aside from algebraic and logical operators a few research work derived the optimum fusion rules based on the assumptions satisfy the output distribution [7, 68, 69]. Majority Voting and Weighted Majority Voting are common techniques to fuse classifier’s outputs. Majority voting makes the final decision by counting the votes and compared them to 50%+1 of total voters. However, to utilize majority voting, assuming all classifiers perform equally accurate must be satisfied which is not true in all applications. The theoretical analysis of number of matchers, their accuracies and the upper boundary of majority voting limit is discussed in details in [70]. Note that, no prior information of the classifiers is taken into account in majority voting and the prior knowledge of training data distribution is neglected in the algorithm.

Figure 2.11: (a) Distributed detection system with data fusion center. (b) Optimum data fusion structure adapted from [7]

- Optimal Discrete Fusion rules

The optimal solution for decision making based on the *a priori* information is to employ Chair-Varshney rule proposed in [7]. In a set of sensors shown in Figure 2.11 (a), P_0 and P_1 are defined as the a priori probabilities of the two hypotheses of both classes. They assumed that there are 'n' detectors and the feature vector or the source at each classifier or detector are denoted by y_i , $i = 1, \dots, n$. Assuming observations at the individual detectors are statistically independent, the conditional probability density function is denoted by $p(y_j|H_j)$, $i = 1, \dots, n$, $j = 1, 2$. Each detector employs a classification algorithm as $g_i(y_i)$ to make a decision ' u_i '. Furthermore, the probabilities of false alarm and miss of each detector (classifier) are denoted by P_F , and P_M , respectively. They derive the final decision making rule of the ensemble of sensors function $f(u_1, \dots, u_n)$ as,

$$f(u_1, \dots, u_n) = \begin{cases} 1, & \text{if } a_0 + \sum_{i=1}^n a_i u_i > 0 \\ -1, & \text{Otherwise} \end{cases} \quad (2-13)$$

where,

$$\begin{cases} a_0 = \log\left(\frac{P_1}{P_0}\right) \\ a_i = \log\left(\frac{1 - P_{M_i}}{P_{F_i}}\right) & \text{if } u_i = +1 \\ a_i = \log\left(\frac{1 - P_{F_i}}{P_{M_i}}\right) & \text{if } u_i = -1 \end{cases} \quad (2-14)$$

The threshold “ a_0 ” is calculated using prior probabilities of two classes which are almost unknown for classification, so this threshold can be set by the training set.

2.10.3 Continuous Score Fusion Techniques

The confidence level or continuous score is the measurement of matching between the query sample and the genuine biometric data template. The ease of obtaining match score and its embedded wealth of information make it an appropriate stage for information fusion as commonly used in multi-biometric systems. Note that, curse of dimensionality, homogeneity, distribution, nature of the score and normalization algorithms are challenges of a perfect score fusion rule. Among fusion methods at different stages of recognition, score fusion is usually preferred due to the balance between information possession and the fusion complexity [48]. It should be well cared if the confidence level rendered by the matchers may have different nature such as similarity measurement or distance that leads to diverse Probability Distribution Functions (PDF).

We assume that there are L matchers with the output similarity score vector (match scores) denoted as $S = [S_1, \dots, S_L]$, to verify if the query sample X belongs to the genuine class or an impostor. The modified Bayesian rule for a system of equal loss of Positive or Negative error rate is,

$$X \in \text{Genuine} \quad \text{if} \quad \frac{P(X \in \text{genuine} | \vec{S})}{P(X \in \text{imposter} | \vec{S})} \geq \eta \quad (2-15)$$

Based on the Bayes rule, the posterior probability can be written as,

$$P(X \in \textit{genuine} | \vec{S}) = \frac{P(\vec{S} | X \in \textit{genuine})}{P(\vec{S})} \quad (2-16)$$

$$P(X \in \textit{imposter} | \vec{S}) = \frac{P(\vec{S} | X \in \textit{imposter})}{P(\vec{S})} \quad (2-17)$$

Therefore, the decision can be made,

$$X \in \textit{Genuine} \quad \textit{if} \quad \frac{P(\vec{S} | X \in \textit{genuine})}{P(\vec{S} | X \in \textit{imposter})} \geq \eta \quad (2-18)$$

In a biometric system the output distribution is not known due to the very large and complex subspaces of biometric traits. Hence, the posterior probabilities in equation (2-15) need to be estimated. Various techniques are proposed in the literature for the posterior distribution estimation as following,

- Transformation-based methods

It is challenging to achieve accurate estimation of the joint conditional densities due to the hassle and the time consuming data collection task. One way is to transform different match scores into a common space to combine them using algebraic/logical operations [71]. It is also known as match score normalization. Transformation based algorithms convert match scores to a normalized common domain and require large amount of data for evaluation.

- Classifier-based methods

In the classifier-based approach, the output scores are concatenated to form a final feature vector and a second stage of classification is performed to obtain the final results of being an imposter or a genuine sample. Based on the training set, the

classifier learns the boundaries of the genuine and imposter samples. The curse of dimensionality does not apply in this method since the classifier easily learns boundaries of different features in nature and dimension. One of the shortcomings of the classifier-based score fusion methods is minimization of an error type while keeping the other constant or to have different losses for each error type. For example, in case of finding missing children which is very critical to fail to spot a genuine sample (minimize P_{miss}), it is complicated to include this criteria in classification based algorithms.

- Density-based methods

Let S_g and S_i be the confidence level random variables of the genuine and imposter samples, respectively. The distribution function of S_g , $F_g(s)$ and the corresponding density function $f_g(s)$ are calculated as,

$$P(S_g < s) = F_g(s) = \int_{-\infty}^s f_g(t) dt \quad (2-19)$$

and correspondingly for S_i ,

$$P(S_i < s) = F_i(s) = \int_{-\infty}^s f_i(t) dt \quad (2-20)$$

The probability density functions $f_g(s)$ and $f_i(s)$ are the short form of conditional densities as

$$f_g(s) = f(s|X \in \text{Genuine}) , f_i(s) = f(s|X \in \text{Imposter}) \quad (2-21)$$

The density estimation is broadly divided to parametric or non-parametric algorithms [72]. The general form of density function is assumed for the distribution in parametric density estimation techniques and parameters are estimated e.g. mean and variance to be estimated for a normal distribution. Non-parametric estimation techniques require large amount of data to estimate the

density functions accurately which is not always available in biometric applications, especially from the genuine class.

2.11 Analysis of Family Members' Faces Data

There are many research work investigating that individuals are able to recognize their kin and behave accordingly to their family members in social species such as ants, insects, monkeys and fishes [73-75]. "Apparent strategies in evolution that favor the reproductive success of an organism's relatives" is referred to kin selection based on Hamilton's inclusive fitness theory [76]. Rejecting new transplanted organs by patients' bodies or spitting out a small tadpole when his own parent engulfs it are the inherent responses of recognizing kin members [77]. Behaviorally distinguishing between individuals with different distance in terms of relatedness requires organisms to be able to distinct cues of kinship.

Humans compare facial information to recognize their relatives and regulate their behavior accordingly [78, 79]. Various research work imply that humans use facial resemblance as a cue to identify their relatives such as making decisions of paternal investment to people with self-resembling faces [80] or attractiveness of the less resembled opposite sex faces [81]. In crime situations, cases of men committed spouse/child abuse are highly proportional to how much they were told that the children do not resemble them [82]. All these emphasize the idea that facial resemblance is quite related to verification of family members from others and affects humans' behavior. The ability to process facial resemblance in human brain is carried out by combination of existing neurocomputational architecture [83]. However, the cognitive processes associated analysis of facial resemblance and family verification algorithms "have received little attention to date" [78].

In the field of computer vision, there are a number of commercial albuming software products, web pages, social communities that provide album organization to the user. The family photo images present a difficult problem for face identification, photo organization and face tagging because of the large differences in appearance due to lighting conditions, pose, image quality and the aging effect. Among the first research papers, M. Das et al. developed a system for automatic albuming of consumer photographs including snapshots from family photo collections. It provides the user with the option of selecting image groups based on the people present in them. Their soft biometric core components of age/gender and facial similarity clustering were utilized to overcome variations of a face during ages, poses and other changes [31].

The age-based information is combined with event and scene segmentation to add more capabilities to the designated system. Their system was able to organize different occasions based on existing people their age and gender like, "Jane's visit to the Museum". The face recognition system employed Euclidian distance measurement of the faces with different clusters of people in the album as the measurement of the similarity [31],

$$d_{min}(C_i, C_j) = \min_{x \in C_i, y \in C_j} \|x - y\| \quad (2-22)$$

where C_i and C_j are two different clusters. The minimum Euclidian distance of clusters to the query image is chosen as the closest or the most similar face. One of the disadvantages is that similar faces of family members were classified as the query person. However, there was no attempt to classify members of a family as a whole class.

The other related research to family verification using faces in family digital albums was done by Singla P. et. al. They investigated Markov logic to "discover social relationships in consumer photo collections" [33]. In addition to the hard rules in

social relationships in family photos like, parents are older than children and soft rules like couples are of the opposite gender or “a young adult occurring solely with a child shares the parent/child relationship”, Markov logic is able to learn new rules using the data. Thereby it enhances the background knowledge by assigning weights to each rule. On the other hand, relationships prediction allowed them to guess the age and gender to improve the accuracy. Although they accomplished guessing social relationships in their collected data but their algorithm is just applicable to family photo datasets and more specifically without any other friends or strangers. On the other hand if in an occasion close relatives like nephews, aunts or uncles swap their position the algorithm fails to discover the correct relationship.

The only related work to family verification was to employ holistic local features of the face [12]. Like the primitive face recognition methods by Kanade in 1973 [24] they extracted various features such as eyes positions, skin color, left and right eye color and etc. They employed the K-Nearest-Neighbor (KNN) by measuring the 2nd order distance of the features to find the nearest class. They concluded that the best feature for family (kinship) verification is the left eye gray-scale color with average accuracy of 72% on the collected dataset of size 286 samples.

2.12 Concluding Remarks

In this chapter, we introduced the state-of-the-art in the domain of facial data analysis. We analyze the most well-known approaches from the holistic appearance based algorithms to highly accurate statistical learning methods. Ordinary modules of the learning algorithms including feature operators, feature selection methods and matching algorithms were studied in details in order to present our contributions which will improve existing solutions. In particular, Section 2.3, 2.4, 2.7.1 and 2.7.2 emphasize the main drawbacks of the conventional face data analysis methods for family data analysis and Section 2.9 and 2.10 focus on finding possible solutions to

increase the accuracy of the state-of-the-art learning algorithms. In the following chapters, we propose the new challenging problem of family verification and propose multiple frameworks to recognize family members.

Chapter 3

Our Proposed Framework for Family Verification and Resemblance Extraction

3.1 Introduction

In this chapter, we explain a classification framework for family verification. The proposed method, a family facial patch resemblance extraction framework ranks facial segments in terms of the amount of similarity between family members. In face recognition, we encounter controlled conditions as various face datasets contain samples of variable photometric changes and camera positions relative to the face [13, 84]. It was shown how to select features required to distinguish between face and non-faces [28]. These features are possessed by all the faces but not in non-face samples. Later on algorithms were proposed to differentiate large group of people as gender. Our aim is to choose common features of members within a family that differ from other families. These common features among a family of human beings

separate them from other group of people that was missing in this chain. Finally, face recognition the last stage of this chain, finds discriminative features that distinct an individual from all the other subjects. Current state-of-the-art face recognition algorithms employ feature selection and matching algorithms to the nonlinear convex face manifolds to learn the decision boundaries [18].

On the other hand, family verification is evaluated on family photo albums collected from various sources, camera sensors, photometric changes, accessories and slight expression changes. Hence, these changes are uncontrolled and inevitable. In family verification, we also encounter larger intra class variability due to multiple individuals from the same class, multi-age groups and genders in the same database. There are currently large face recognition databases available that eases the performance evaluation for researchers [2]. We collected a database of 45 family albums from volunteers, friends and available online shared albums. The database contains more than 5400 face samples including families of different ethnicities, with high aging effect in the parent's face images and other variations may occur in human families. To evaluate family verification using various algorithms different scenarios were defined to cover real life applications. The challenging problem of unseen members is also investigated to contribute to finding missing children as a critical application of family verification. Mathematical modeling of the family verification problem reveals that the conventional face recognition algorithms need to be improved to find minor similarities of family face members. We could enhance the performance using the mixture of experts since insufficient data, data types and "divide and conquer" strategy are satisfied in human families.

In section 3.2 to 3.5 we study principles, challenges and assumptions of family verification. The whole problem is generally defined to cover demanding applications in computer vision. We consider famous face recognition methods and evaluate their performance on family verification problem in section 3.6. Results of face recognition methods in section 3.6 expose the essential need for development of a framework for

family verification. Finally, we propose a general framework to recognize family members using the facial resemblance among them regardless of feature operators and matching algorithms in section 3.9. The family members' faces are split to regions. Features are extracted from each patch and passed through the matching stage. The family verification reliability of each patch ranks the similar patches in each family dataset. The Chair-Varshney fusion rule finds the final decision using the provided decision from each individual patch. We analyze the information generated from each patch to select the redundancy of the facial region for family verification. The inter-patch, constraint-free analysis of the face segments reduces number of required patches for family verification. This chapter ends with experimental results and discussions on family datasets and illustration of the family resemblance for selected families. The computational saving achieved by a limited number of patches is also examined and tabulated to highlight its supremacy for family verification.

3.2 Challenges of Family Verification

Researchers benefit from the face data information to enhance human identification performance aside from using the face image solely for face recognition [21]. It has been rarely looked into recognizing members of the family as a whole class to verify their relationship [12]. The genetic resemblance of family members to one another, especially among siblings is the key point to verify the relationship of a person and his/her immediate family members. We investigate automatic family member verification among human family datasets to enable the artificial intelligence of this inherent human capability.

In almost all face recognition applications, the individuals to be recognized have been enrolled *in advance*. Thus, there is at least one sample in the training dataset. The challenging problem of one genuine sample face recognition has attracted the researchers to come up with strong algorithms to tackle this problem [85]. We intend

to push the boundary and consider a more demanding problem where we endeavor to recognize the person without any sample and utilize his/her family samples for recognition. There are interesting scenarios where it is imperative to automatically identify possible candidates of a specific family member whose photo is not available. On average, 2,185 children are being reported missing each day, such as at megamalls. It would be helpful to be able to find the missing child in a surveillance camera system by utilizing a few samples of their parents and other children [23]. The problem of family verification differs from the conventional face recognition in various aspects as following,

- Intra-class Variability

Family verification has larger intra-class variation compared to face recognition. The algorithm must be able to generalize the recognition capability such that it can recognize different faces that possess similar facial features that characterize members of a family to differentiate from those similar faces but not having those facial features. In a face recognition system, we search for exact features of the face in the query sample. However, family resemblance may not appear as the exact features in members' faces due to genetic proximities. These similar features must be extracted from face variations due to different age and gender even if lighting, pose and other variations are minimized. As such, for family verification, the similarity of the query face is compared with multiple faces from different identities having similarities which are common among all family members. If similar features are found in the query image, it is classified as a member of the family. The main problem in family verification is that the family members' intra-class similarities are much less than face recognition since there are only limited features to characterize the common family facial features. Figure 3.1 shows a series of sample images of President Obama's face at the top row and his family's images at the bottom row. These images show the contrast

between individual and family face characteristics. It can be seen that the resemblance and the discriminant features in the 1st row (face recognition) are notably more than the family samples in the 2nd row (family verification). Therefore, more discriminate features are required to extract the available resemblance information.

Figure 3.1: Samples of face recognition and family verification for President Obama (first row) and his family members' face samples (second row)

- A Rather Complex Face Subspace

The face subspace is a complex convex manifold due to image variations. There are other challenges in family verification which are different from face recognition. Face recognition aims to find a matching face in a bank of samples (or database) given a query face. Generalization of face recognition is only required to take into account variations due to face pose, facial expression, face make-up or facial hair, wearing of spectacles and lighting. If the template is not refreshed regularly, then the face recognition algorithm will also have to take into account facial variation due to aging. The family datasets encounter the inevitable changes of members' genders, age difference of family members and possible mixed ethnicity parents. Family verification requires tackling all

variations found in face recognition and those inherit changes of family members. Moreover, unbalanced datasets (i.e. some family members may have more photos than the other), mixed ethnicities and etc. are additional variations encountered in family verification but not in face recognition. For children, twins or quadruplets may increase the performance of recognizing other siblings but also provide less divers information of the whole family.

The first step of analyzing a pattern classification problem is the data collection. There are many face databases in use the choice of an appropriate database depending on what circumstances i.e. aging, expressions or lighting [13] the problem is going to be solved [2]. Samples of a face recognition dataset (controlled conditions) are depicted in Figure 3.2 (a). It is relatively easy to be processed compared to family albums depicted in Figure 3.2 (b) with inevitable changes of ethnicity, age, background and accessory [16]. Such variations increase complexity of face subspace in family databases.

3.3 Family Dataset and Scenarios

Unlike face recognition, there is no publicly available dataset for family verification comprising photos of family members taken over a period of time specifically for family verification except the Gallagher's dataset [16]. For our experiments, we collected a family database comprising 45 families with an average of 120 face samples per family (overall 5400 samples) from the on-line digital albums, friends and volunteers. Thus, photometric changes such as lighting, background, accessory, camera lens and sensor variation inevitably exist in collected digital albums. The dataset contains families of different ethnicities.

Figure 3.2: Face database samples (a) face recognition CAS-PEAL database adapted from [13] (b) Family albums samples adapted from [16]

It includes families with high aging effect in the parent's face images, families with adolescences, 3 families with twins or quadruplets and 6 families with single child. All faces are cropped with respect to their eyes' position extracted using the Active

Appearance Model (AAM) [25]. As our current work does not consider significant pose variation, only faces with near frontal pose are used. Finally all face samples are resized to 80×95 pixels and based on the experiment required down sampling is easily performed. Recently, a research group published a database named “kinship database” from Cornell University that includes 143 individuals and their corresponding parents with just a single sample per person. The average face size is about 50 pixels and the main shortcoming of the dataset is only a single image is provided for each family member that restricts algorithms such as statistical learning methods to be utilized [12]. The variety of family specifications and number of families with characteristics of our database are tabulated in Table 3.1. Diverse samples of the family database are depicted in a thumbnail album in Figure 3.3. It is obvious in Figure 3.3 that variations such as occlusion, accessories, background, lighting slight pose and expression is inevitable in collected photos.

3.4 Experimental Scenarios and Performance Evaluation

The aim of this thesis is to propose a general framework for family verification to suit applications and critical situations around us. In order to conduct challenging experiments to compare our proposed approach with the state-of-the-art face recognition algorithms, we define 3 different scenarios of family verification to explore the possible situations in the real world:

Test 1: There are situations that family members are present in both training and test set, as in the application of family verification for digital cameras and human computer interaction.

Table 3.1: Family specifications and their characteristics in our collected database

Characteristic	Quantity
High Aging effect (max Age difference between parents and children Parents is 40)	18 families
Medium Aging effect (max Age difference between parents and children Parents is 30)	19 families
Low Aging effect (max Age difference between parents and children Parents is 30, politicians' or artists' families)	8 families
Asian (Chinese), Parents' ethnicity	11 families
African – American, Parents' ethnicity	6 families
Caucasian, Parents' ethnicity	26 families
Mixed Ethnics, Parents' ethnicity	2 families
Siblings are twins, triplets or quadruplets	3 families
Families with children of adolescence age group	16 families
Number of children in families respectively [1,2,3,4]	[7,31,5,2]
Album extracted from move frames	4 families
Taken by high quality camera	23 families
Parents/siblings wearing glasses	9 families
The album time span [1,2,3,more than 5] years	[21, 9, 4, 11]
Average number of samples for [parents, children]	[12, 27]
Families with related parents (to our knowledge)	None

Figure 3.3: Samples of 4 selected family datasets with inevitable changes of occlusion, accessories, closed eyes, background, lighting, and slight pose and expression changes. Parent's samples are shown at the first row and their corresponding children at the 2nd row.

Test 2: In some situations a member may not be present or known i.e. missing child or unknown father. Two different scenarios are defined:

Test 2.a in which the sibling with max number of samples is omitted from the training set

Test 2.b in which one of the parents with max number of samples is omitted from the training set.

The family member with the largest number of samples is chosen as the omitted member as this gives the largest error to make the test more challenging. Moreover, we obtain more comprehensive error evaluation. In test 2, all family members' samples, except the omitted member, are used for training together with a portion of nonfamily members. For the test set, only the omitted member's samples exist together with the remaining nonfamily members. The performance is evaluated by considering three types of errors as:

- False Negative Ratio, FNR, denoted as samples of members of the family that are not recognized as family divided by the whole number of samples of the family members in the dataset.
- False Positive Ratio, FPR, denoted as samples of non-family members that are recognized as family divided by the whole number of non-family samples in the dataset.
- Total Error rate, Err, denoted as all misclassified samples divided by total number of samples

Our goal is to minimize the FNR to reduce family member misclassification.

3.5 Family Verification Modeling

Face identification problem is as a two-class problem. The genuine class contains face samples of the individual I with the $\bar{X}_I$ set comprising images of individual I

$$\bar{X}_I : \{X^1, X^2, \dots, X^d\} \text{ belongs to Individual } (I) \quad (3-1)$$

In a face identification problem we classify the query image (X_Q) to class w_I genuine data or w_0 imposter data. The similarity measurement method $S(X_Q, \bar{X}_I)$ determines the maximum resemblance of the given sample with the bank of genuine template. Finally the decision is made by,

$$\begin{cases} w_1 & \text{if } S(X_Q, \bar{X}_I) > \eta \\ w_0 & \text{otherwise} \end{cases} \quad (3-2)$$

where η is a preset threshold. In family verification, we have multiple members and we aim to determine if the query sample is similar to all members. The assumption of belonging the sample to one of the members is not valid since we may encounter missing member or the provided samples may have acquired at very long time span. We assume that family (i) denoted as F_i has m members including $\{M_1^i, M_2^i, \dots, M_m^i\}$. Hence, the family verification is formulated as [2],

$$\begin{cases} C_1 & \text{if } S(X_Q, \bar{X}_{M_1^i}, \dots, \bar{X}_{M_m^i}) > \eta \\ C_0 & \text{otherwise} \end{cases} \quad (3-3)$$

where η is a preset threshold and $\bar{X}_{M_j}$ is the set of face samples belong to member (j) of family (i) denoted as M_j^i . X_Q is the query face sample. The genuine class of family members is denoted as C_I and belonging to others (nonfamily samples) is denoted as C_0 . The similarity score measurement is a matching algorithm that learns the complex face boundaries of the selected discriminative features. The manifold of face

samples from individual I is denoted as $\text{MF}(\bar{X}_I)$ and consequently for M_j^i denoted as $\text{MF}(\bar{X}_{M_j^i})$. Based on genetic proximities,

$$\bigcap_{j=1}^m \text{MF}(\bar{X}_{M_j^i}) \neq \emptyset \quad (3-4)$$

and for every member (if j & k are not parents),

$$\text{MF}(\bar{X}_{M_j^i}) \cap \text{MF}(\bar{X}_{M_k^i}) \neq \emptyset \quad \forall j, k \leq m \quad (3-5)$$

Based on the intersection rules the intersection of the query face's manifold with all family members is the subset of intersection of the query face's manifold with each member from equation (3-6)

$$\begin{aligned} \bigcap_{j=1}^m \text{MF}(\bar{X}_{M_j^i}) \cap \text{MF}(X_Q) \\ \subseteq \text{MF}(\bar{X}_{M_k^i}) \cap \text{MF}(X_Q) \quad \forall j, k \leq m \end{aligned} \quad (3-6)$$

It shows that the corresponding number of discriminative features of the query and the genuine face manifolds intersection in family verification is smaller than that of face recognition based on,

$$\text{if } A \subseteq B \text{ then } |A| \leq |B| \quad (3-7)$$

where the cardinality ' $| \cdot |$ ' of a set is the number of members of that set. Similar to extension of face identification to face recognition the task of family verification can be extended to family recognition as,

$$\begin{cases} F_i & \text{if } i = \arg \max_i \{S(X_Q, \bar{X}_{M_1^i}, \dots, \bar{X}_{M_m^i})\} \text{ and } S(X_Q, \bar{X}_{M_1^i}, \dots, \bar{X}_{M_m^i}) > \eta \\ \text{Non-family} & \text{otherwise} \end{cases} \quad (3-8)$$

It is also possible to measure the distance using Mahalanobis n -th order to check the dissimilarity instead of resemblance to the family members. We continue this chapter by proposing a novel framework for family verification to extract family facial resemblance as the intersection of members' facial manifolds.

3.6 Family Verification Using Conventional Face Verification Approaches

Face recognition is the closest work to explore family verification. Nowadays, many successful algorithms for analyzing face data are learning-based [2]. As aforementioned in section 3.5, finding the intersection of family members' face manifolds makes the family class boundaries highly non-linear. Popular approaches to overcome the nonlinearity employ local appearance-based feature spaces. The overall block diagram of the conventional face recognition methods is depicted in Figure 3.4.

Figure 3.4: General block diagram of conventional face recognition methods

Firstly, we analyze feature selector and classifier algorithms of conventional face recognition methods and their performance on family verification. The first stage is the feature selection method. As discussed before, linear and holistic approaches are not accurate enough to find non-linear boundaries of face manifolds in the face space. AdaBoost [8] is a widely used feature selection and learning algorithm that combines

many weak classifiers to form a strong classifier. It adjusts the weights of training samples given their classified labels compared to the actual true labels iteratively. AdaBoost has been used successfully to handle the feature selection and nonlinear classification problems [2, 6]. On the other hand, kernel methods may not perform well on the unseen data [2] which occurs in family verification. The K-Nearest Neighbors (KNN) and the Support Vector Machines (SVM) algorithms are very well-known in pattern recognition applications [72]. SVM is currently among the strong classifiers for face verification and fusion techniques[2, 72]. We explain these two classifiers in summary.

3.6.1 K-Nearest Neighbors

The K-Nearest Neighbors algorithm (KNN) is an instance based method for classifying samples based on a majority vote by the k training examples closest to a query sample in the feature space [86]. Mahalanobis distance measurement could be utilized as dissimilarity score calculation. Euclidean distance is the default distance measurement used in this algorithm. However, one needs to adjust the parameter K to obtain the best performance. The training data may contain outliers such as samples taken in very poor lighting conditions or extreme pose angle. The strangeness measurement as the basis of Transduction proposed in [87] could be also investigated to remove dissimilar, unlikely and far genuine sample to prevent highly nonlinear boundaries. One can benefit from less strange samples to train the KNN classifier.

The ratio of the sum of the k-nearest neighbours from the same class y divided by the sum of the k nearest neighbours from others ($\neg y$) is defined as strangeness. So the strangeness of i_{th} sample is defined by α ,

$$\alpha_i = \frac{\sum_{j=1}^k d_{ij}^y}{\sum_{j=1}^k d_{ij}^{\neg y}} \quad (3-9)$$

3.6.2 Support Vector Machines

The Support Vector Machines (SVM) is a supervised learning method that classifies two data sets by constructing the separating hyper plane using kernel functions. The separating hyper-plane distance is farthest from the data points that maximizes the margin. This would prevent from over-fitting and minimizes the risk of error when classifying new sets of data. The resultant hyper plane is designed to maximize the margin between these sets in feature space, to give the best possible separation between the training samples from both classes, according to some pre-defined penalty metric as shown in Figure 3.5 [1, 88].

Figure 3.5: SVM Optimal Separating Hyper-plane adapted from [1]

There are mainly two types of SVM classification, linear and non-linear. The less complicated hyper-plane, linear kernel classification, separates the data into different classes. The kernel function transfers the nonlinear data to a higher dimension space. The data is then transformed to a “feature space” to obtain linear separation [72]. The Radial Basis Function (RBF) kernel is mostly used for challenging classification and fusion techniques [89]. The γ parameter needs to be adjusted for RBF kernels to obtain the best classification result as stated in the RBF kernel equation,

$$K(x_i, x_j) = \exp\left(-\gamma \|x_i - x_j\|^2\right) , \quad \gamma > 0 \quad (3-10)$$

Currently, there are available sources to implement SVM classifiers providing methods to set parameters in MATLAB as [1, 90].

3.7 Benefits of Mixture of Experts for Family Verification

As aforementioned in Chapter 2, the general idea of consulting multiple experts for enhancing classification performance is widely investigated [46]. Family verification problem meets the conditions aforementioned to benefit from MOE for performance improvement. In this section, we discuss the possible contributions of different MOE strategies to recognize family members based on the realities and assumptions of family members' facial resemblance.

- Statistical Reasons

Family members must be representative of the main family face manifold so that they cover all variations and similarities of the high dimensional family data. In other words, specifically in Test 2 when a members (subclass) of the family is absent in training set, we expect MOE to improve the classification performance more than other tests by obtaining more information of the face data distribution.

- Too Little Data

Family datasets digital albums are difficult to be collected and only limited samples of each member are available publically due to privacy. In some cases, previous researchers refuse to publish their collected database over the internet or do not share photos [32]. Family verification that encounters too little samples size can benefit from MOE [46].

- Divide and Conquer

As shown in Figure 2.10, assuming if each elliptic class represents a family member then the MOE will help improve the performance of family verification. Conquering each part of the face or member's patch extracts the partial resemblance. It is more critical as the similarity of family members is affected due to aging effect, different races or age groups in the same album. In this way each classifier becomes an expert to recognize each face segment or member belongs to family.

3.8 Data Usage for Experiments

To the best of our knowledge, family verification as generally defined in this research work was not analyzed before in the field of computer vision. Moreover, there is no discussion in face recognition research work [18] solely investigating family albums to recognize members of the same immediate family relationship. Therefore, it is essential to examine conventional face recognition approaches and understand their advantages and shortcomings to tackle family verification challenges. We initiate evaluation of classifiers on family verification based on the experiments defined in section 3.4.

- Data Preparation

The faces of the adult family members are cropped with respect to the eyes' positions [25]. The two straight face samples of 554 individuals of the first DVD of the FERET dataset [84] are cropped with respect to the provided eyes positions. 300 random samples from FERET database and 800 nonfamily members from family albums create the non family database. For each family, 2/3 of the family dataset and 2/3 of the non-family dataset are used for training and the remaining are used for testing. The experiments are conducted for test 1, 2a and 2b. For test 2a and 2b, 2/3 of all family members' samples, except the omitted member, are used for training together with 2/3 of nonfamily members. For the test set, all samples of the omitted member and the remaining non-family members are used. The classification problem is to distinguish between samples belonging to Family #i as class "1" denoted as C_1 and nonfamily members as class C_0 .

In our study on family verification and exploration on distinctive features from conventional feature extraction methods publication [8] in section "Author Publications", we noticed that Gabor features are well capable of extracting facial resemblance [40]. The total error of Gabor wavelets and the AdaBoost classifier was the least between selected features. Hence, we adjust parameters of Gabor wavelet banks with 5 scales and 8 orientations to extract features of the cropped face images (as samples are depicted in Figure 3.6). The face samples were cropped to size 20×24 for simulations. AdaBoost is selected as the feature selector as in Figure 3.4. For classifiers, we selected AdaBoost classifier, SVM and KNN. As in the literature, extracted Gabor features were passed through the AdaBoost feature selector [6]. Dominant features picked by the AdaBoost after 200 iterations are given to the classifiers to find the decision boundaries [6]. We employ the strong SVM classifier with Radial Basis Functions.

Figure 3.6: Samples of cropped faces with respect to eyes positions of families 1 to 3 at each row, parent's samples are the first two samples and the rest of photos are their corresponding children

To implement SVM, firstly, the input data are scaled to the $[-1, +1]$ range for each feature. Then we employ five-fold cross-validation to determine the most optimal values for the SVM parameter values C , the penalty term and γ , a crucial radial basis kernel parameter for this data set. The SVM classifier is evaluated at various coarse combinations of these two parameters, in order to identify the approximate range of optimal parameter values for further parameter fine-tuning. The final set of parameter values are $C = 2.0$ and $\gamma = 1.5$. The KNN classifier is also in need of parameter adjustment, the number of neighbors to choose for distance measurement. The performance is evaluated for K between 2 to 8 which is 2 is chosen for each family. It is interesting that the performance degrades for K greater than the number of family members present in the training set. The error rate of selected classifiers and LDA [52] are tabulated in Table 3.2.

Table 3.2: AdaBoost, SVM, KNN and LDA average performances (FNR, FPR, Err) for different tests

Classifier	Test 1			Test 2.a			Test 2.b		
	FNR	FPR	Err	FNR	FPR	Err	FNR	FPR	Err
AdaBoost	0.267	0.010	0.04	0.34	0.008	0.022	0.52	0.007	0.015
SVM	0.09	0.03	0.042	0.12	0.023	0.029	0.12	0.086	0.085
KNN	0.21	0.034	0.053	0.335	0.027	0.04	0.47	0.02	0.029
LDA	0.65	0.083	0.12	0.77	0.073	0.1	0.83	0.062	0.092

The results show that SVM performs better than other classifiers in terms of FNR. However, the total error of AdaBoost is much less than KNN and SVM and that is due to very low FPR achieved by the AdaBoost classifier. As AdaBoost has the most accurate results in terms of total error and the FNR achieved by this method is higher than SVM we try to minimize its FNR when keeping FPR constant to improve its accuracy.

We also study the effect of using Majority Voting (MV) (i) on the outputs of these classifiers and (ii) if an individual classifier is trained for each member. If the output from the classifiers shows a majority for the candidate to be a family member, then the sample is considered as family, otherwise as nonfamily. For cases where there are even number of classifiers (ex: family 1 and 2 in test 1), the selected threshold is equal or more than half the total number of available classifiers. This was done to reduce the false rejection rate (where a family member is classified as non-member) which is a more critical error for family verification. Table 3.3 shows the results of MV on the outputs of multiple AdaBoost classifiers on each individual member in three selected iterations to check if computational cost could be saved by such fusion. The results reveal that utilizing face recognition on each individual family member does not necessarily lead to higher accuracy for family verification.

Table 3.3: Majority voting on the outputs of AdaBoost classifiers for each individual family member for different tests in selected iterations

Iteration No.	Test 1		Test 2.a		Test 2.b	
	FNR	FPR	FNR	FPR	FNR	FPR
5	0.42	0.051	0.386	0.058	0.613	0.052
50	0.386	0.035	0.373	0.048	0.496	0.05
200	0.386	0.035	0.373	0.048	0.496	0.05

An interesting observation in using MOE on individual members is that family verification performance improves slightly for unseen family member if family members belong to different age groups.

The other way to benefit from MOE is to use mixture of selected classifiers AdaBoost, SVM and KNN, implemented before. Therefore Majority Voting (MV) is implemented on the outputs of these three classifiers to increase the performance. The results are shown in Table 3.4 for AdaBoost, SVM and KNN on 600 selected features by AdaBoost after 200 iterations.

Table 3.4: Majority voting performance on the outputs of multiple classifiers for different tests

Classifier	Test 1		Test 2.a		Test 2.b	
	FNR	FPR	FNR	FPR	FNR	FPR
MV on selected classifiers	0.143	0.0113	0.236	0.0093	0.236	0.032

- Minimization of FNR

The aim of family verification is to recognize family members which is equivalent to minimization of FNR. Among the classifiers chosen in section 3.8, AdaBoost outperformed other matching algorithm. However, it requires design of fusion or final threshold selection method to minimize the FNR while keeping the FPR at constant level.

- Resemblance of Family Members

The key solution to the family verification problem is to extract facial resemblance among members. The experiment of utilizing an individual classifier for each member underperforms to recognize family members compared to considering all members as a whole class. Moreover, finding similar faces is generally possible if only two faces are to be examined [36]. The resemblance among family members exists in different parts of the face and differs from family to family due to genetic proximities. Hence, the algorithm of facial resemblance extraction must be able to find general similarity among members and be robust against possible occlusions. More importantly, we need to know the amount of similarity of each facial part and the amount of information embedded in facial parts. Finally the algorithm utilizes the extracted resemblance information to perform family verification.

3.9 Extraction of Family Facial Patch Resemblance

In this section, we improve the family verification performance based on their facial resemblance measurement and remove the redundant facial patches. The proposed

method performs family verification regardless of the classifiers and features operators. Features of split facial patches (Data #i) are extracted and given to individual classifiers. The optimal decision making rule is employed on classifiers outputs. The final block identifies the redundant information and adjusts the weights of each patch using the training data. In this manner, the overall system is trained for the family dataset, to classify family members based on their facial resemblance measurement and remove redundant information. The overall approach is as follows:

1. Split facial patches using the modified face template
2. Features of each patch are extracted and discriminative features of each patch feature set are selected by AdaBoost
3. The most informative set of patches are selected for recognition and the other patches are removed to give a minimal set of patches for recognizing the family.
4. Based on the reliability of each patch, the optimal Chair-Varshney decision making algorithm is implemented for family verification

The block diagram in Figure 3.7 illustrates the whole algorithm. Detailed explanation of the proposed algorithm is given in the following sub-sections.

Figure 3.7: The overall framework of family verification using informative facial resemblance

3.9.1 Facial Patch Template

The distances of the faces' interest points have almost the same ratio, known as the Golden Ratio [91]. This average ratio varies for different ethnicities and also from face to face. Since the members of a family have some resemblance among them, we postulate that the ratio would not vary for facial patches if the size of the face image is normalized. The modified face template obtained by Marquardt [37] does not include facial similarities of family members such as chin and face shape and the whole region of eyes or nose. To maintain ratios of the modified golden ratio mask and include similar family facial patches, the face template and its corresponding indexing as depicted in Figure 3.8 (b) is proposed. Figure 3.8 (a) also shows four members of family 3 with their split patches.

Figure 3.8: (a) The proposed facial template and cropped facial patches (b) patch indexing

3.9.2 Resemblance Extraction

The face is split to facial parts (Data #i) and the feature extractor provides the data for feature selector and matching algorithm as depicted in Figure 3.7. The word “similarity” is defined as “an aspect, trait, or feature like or resembling another or another's” [90]. The goal of family verification is to find the common features

(extracted from the face) in the given image and family members' templates using the matching algorithm. Hence, the performance of each individual classifier is proportionally related to the amount of similarity of family members in that patch. Note that the alterations and photometric changes may degrade the matching performance of surrounding patches and lower their rank in terms of the resemblance amount. The performance of the facial patch matchers will reveal the amount of resemblance among a family dataset. During the training stage the patch resemblance ranks are revealed. The next step is to make the final decision based on the obtained results for the query image. The computational load would benefit from analysis of the provided resemblance information by each patch. In the next section we select the most informative patches to remove those do not contribute or carry the information captured previously.

3.9.3 Decision Making Rule

To obtain the final decision based on the outcomes of various patches, we require the combination rule. There are many types of decision making rules used in the literature such as Majority Voting, (MV), Weighted Majority Voting, (WMV) or algebraic operators such as minimum, maximum, median, product or sum rule. The optimal solution for decision making based on a priori information is to employ the Chair-Varshney rule proposed in [7]. The final decision ' f ' is given by:

$$f = \log\left(\frac{P_1}{P_0}\right) + \sum_{S_1} \log \frac{1 - \text{FNR}_i}{\text{FPR}_i} + \sum_{S_0} \log \frac{\text{FNR}_i}{1 - \text{FPR}_i} \quad (3-11)$$

where P_1 is the *a priori* probability of a sample to be a family member (denoted as C_1) with $P_1 = p(C_1)$ and vice versa, P_0 is the *a priori* probability of a sample to be a non-family member (denoted as C_0) with $P_0 = p(C_0)$. There are ' N ' classifiers, here $N = 17$ is chosen since there are 17 patches. The AdaBoost classifier output u_i , $i=1, \dots, N$ will be "1" if it is a family (C_1) and "-1" if it is a nonfamily member (C_0).

$S_1 \in \{u_i \mid C_1\}$ and $S_0 \in \{u_i \mid C_0\}$. FPR_i and FNR_i are the i -th classifier positive and negative error rate respectively. In equation (3-11) all parameters, except the value $(\log(P_1/P_0))$, can be easily calculated using the training data. The value $(\log(P_1/P_0))$ is then set as the threshold in the decision making. Since the goal is to minimize FNR using the nonfamily data, after training the classifiers, based on the Neyman-Pearson criterion, we set the desired FPR indicated by FPR_D as a constant value to obtain the threshold. Hence, the optimum decision making rule with the least FNR is achieved using equation (3-11) for the desired FPR.

3.9.4 Selection of Important Patches

We can measure how reliable each patch is to classify a family member by evaluating the matching algorithm performance on the extracted features from the split patches of the family members' faces. This is possible since each patch has different degree of similarity among the family members and thus their rank for family verification differs. The patch resemblance is thus family specific. After identifying facial patch similarity rank, we proceed to find the informative patches among the split patches of a family. The most informative set of b -patches are the ones which bear close resemblance among the family members as identified by the classifier (' b ' is the number of patches chosen, $b=2, \dots, N$). The set of chosen informative patches, should also be able to generalize to other members who are absent in the training set as an indication of the ability to generalize these informative patches.

In the MOE, diversity of classifiers is the key criterion to obtain good performance. The main idea of MOE is to have a pool of classifiers in the way that each classifier has good performance on a different part of the data samples such that the combination provide the best results for the entire data samples. In other words, they should be as diverse as possible. There are several measurements of diversity in the literature, including entropy measurement or measure of difficulty. The most appropriate approach being capable of diversity calculation for ' b ' out of a total of

‘ N ’ classifiers is to calculate the diversity on any set of classifiers with respect to all other patches to evaluate the set of the most informative patches for a certain family. The diversity measurement of b -selected classifiers out of total N ones, is given by ‘ ρ_b ’ [92],

$$\rho_b = \frac{bN^f}{N_T - N^f - N^t + bN^f} \quad (3-12)$$

where, ‘ N^f ’ is the number of all cases that ‘ b ’ classifiers are all false, ‘ N^t ’ is the number of all cases that ‘ b ’ classifiers are all true, ‘ N_T ’ is the total number of cases (conditions) and ‘ b ’ is the number of selected experts. In order to calculate all classifiers’ diversities or patches’ informativeness, we must find all combinations of ‘ N ’ classifiers’ diversities, which leads to the number of ‘ $N!$ ’ permutations. Diversity is a kind of correlation, since $corr(p,q)=corr(q,p)$ for real outputs of the absolute value of Gabor features matching result, The total number of permutations can be calculated as,

$$Total\ Permutations = \binom{N}{2} + \binom{N}{3} + \dots + \binom{N}{N} = \sum_{i=2}^N 2^i = 2^N - N - 1 \quad (3-13)$$

Diversity of a set of classifiers must be minimum for the best combination of informative classifiers with the least error. Among every set of highly diverse b -patches, the training error is evaluated by the optimum decision making rule and the set of b -patches with the least error is selected as the optimal set for ‘ b ’ number of informative patches. Based on this, it is possible to find the minimum number of informative patches beyond which the performance will not differ much. In the proposed scheme, not only confidence level of patches is considered, but also the amount of diverse information each patch carries, through the diversity analysis. Hence, the whole strategy recognizes family members using the amount of similarity of patches. The computational cost is reduced by measuring the provided

information diversity from the patches to select b-patches from N number of total ones.

3.10 Experimental Results and Discussions

We conduct experiments to extract facial resemblance of the family datasets. Gabor wavelets are selected for feature extraction and the AdaBoost algorithm is employed for feature selection and matching. The data is prepared as in section 3.8 and cropped to size 80×95 to find detailed resemblance features of the face. Test 1, Test 2.a and Test 2.b are conducted with the same data division for training and testing. The chart in Figure 3.9 shows the family verification performance in different test cases while $FPR_D=0.02$. The x-axis is the ascending family index.

Figure 3.9: The Family verification performance for family datasets in defined tests

Table 3.5 shows a sample of each member from family 1-5 and the corresponding normalized facial patch resemblance ranking using the patch reliability measurement color map at the last column for test 1. The results for other families and tests are not shown due to limited space. Note that, the facial resemblance symmetry may be distorted due to hair style, face shape or background.

Table 3.5: Family facial patch resemblance shown by patch reliability measurement color map

#Fa	Father	Mother	Children			Similarity	
F1
F2
F3
F4
F5

The next step is to find the most informative b-patches as described in section 3.9.4. In the literature, the feature subset selection methods assume that the importance of each feature set is independent from the others and the kernels should be located at fiducial points [93]. However, these two constraints lead to sub-optimality of the subset selection. In the proposed method, the importance of each patch is being evaluated with dependency on the rest of the patches to achieve an inter-patch analysis without constraints on the feature subsets. As shown in Figure 3.10 the informative patches found are not only around the fiducial points.

Family verification results obtained for family 3, test 1 are tabulated for both training and testing sets in Table 3.6 with respect to the number of patches (b). The right-most column of Table 3.6 shows the total error (Err) which is obtained by calculating the number of samples classified incorrectly over the *total number* of samples. The FPR_D is set to 0.02 to obtain the threshold value as per equation (3-11) after 25 iterations of training. The results obtained show that even with small b , the total error is only 0.02 to 0.03 more than using all 17 patches. For example, even seven patches ($b=7$) would be enough to achieve almost the same performance as using all patches for family 3 in test 1. The trend of decreasing FNR with increasing number of patches is similar in both training and test sets in Table 3.6. The tabulation for other families and test cases are not included due to limited space.

The summarized results of the minimum number of patches to achieve almost the same accuracy of 17 patches are stated in Table 3.7. The patch indexes and the FNR (test set) achieved with b -patches, is also tabulated to compare with the FNR (test set) of all patches for family verification of families 1-5 and test 1. Note that, patches are sorted based on their index order in Table 3.6 and Table 3.7.

The histogram of patches selected as the most informative ones in Table 3.6 for family 3 is shown in Figure 3.10 (c) to find the importance of each patch in family verification using the color map as shown in Figure 3.10 (f). The same analysis is applied for family 1, 2, 4 and 5 to extract the informativeness measurement of facial patches and the results are shown in Figure 3.10 (a), (b), (d), (e) respectively. This histogram is the measurement of each patch's informativeness for family verification in test 1. However, it does not mean unimportant patches are not reliable but they do not add information to the most informative patches selected. This supports the human intuition of recognizing the family relationship based on judging the similarity of certain parts of the face. Thus by finding such parts, family verification performance can be optimized.

Table 3.6: Determination of b informative patches for family 3 classification, test 1, with the FNR error in both training (Tr) and testing (Te) sets and testing stage total error (Err).

b	FNR(Tr)	FNR(Te)	Patch index	Err
2	0.181	0.481	1, 7	0.10
3	0.152	0.468	7, 14, 16	0.10
4	0.115	0.500	4, 7, 14, 16	0.11
5	0.072	0.451	2, 4, 7, 14, 16	0.10
6	0.059	0.467	1, 2, 4, 7, 14, 16	0.10
7	0.048	0.387	1, 2, 4, 7, 13, 14, 16	0.09
8	0.086	0.387	1, 2, 4, 7, 13, 14, 15, 16	0.09
9	0.072	0.388	1, 2, 4, 7, 8, 13, 14, 15, 16	0.09
10	0.072	0.388	1, 2, 4, 6, 7, 8, 13, 14, 15, 16	0.09
11	0.072	0.396	1, 2, 4, 6, 7, 8, 13, 14, 15, 16, 17	0.09
12	0.072	0.387	1, 2, 4, 6, 7, 8, 11, 13, 14, 15, 16, 17	0.09
13	0.094	0.338	1, 2, 3, 4, 5, 6, 8, 11, 13, 14, 15, 16, 17	0.08
14	0.043	0.388	1, 2, 3, 4, 5, 6, 8, 9, 11, 13, 14, 15, 16, 17	0.09
15	0.057	0.403	1, 2, 3, 4, 5, 6, 7, 8, 9, 11, 13, 14, 15, 16, 17	0.09
16	0.050	0.387	All without patch 12	0.09
17	0.043	0.371	All 17 patches	0.08

Table 3.7: Minimum number of patches needed to achieve almost the accuracy of all patches for family verification of families 1-5 and test 1

#Family	$\min(b)$	Patch index	FNR(Test) $\min(b)$ -patches	FNR(Test) All patches
1	4	1, 2, 3, 4	0.03	0.03
2	4	4, 5, 8, 13	0.22	0.2
3	7	1, 2, 4, 7, 13, 14, 16	0.387	0.371
4	4	1, 2, 3, 7,	0.083	0.08
5	4	1, 2, 3, 4	0.02	0.02

Figure 3.10: The importance histogram color map for family (a) 1, (b) 2, (c) 3, (d) 4, (e) 5 and test 1; (f) color map indexing; (g) feature points localization

To optimize the patch selection, redundant patches has to be removed while keeping the accuracy achievable as high as possible. The constraint-free inter patch analysis of facial patches is proposed to select b-number of the most informative patches. The results obtained show that such selection can provide results close to using all patches. Depending on the family, just using as low as 4 patches provides almost similar performance compared to using all 17 patches.

We also study the effect of background in patch selection for patches 13 and 15. We conduct an experiment by changing the background region for these two patches for 5% and 15% of the images in each family. The results show that patches 13 and 15 are still chosen for those families they were contributing to facial resemblance. Moreover, the selected feature point localization depicted in Figure 3.10 (g) do not include pixels from background but boundaries of face oval.

3.11 Concluding Remarks

This chapter explains our novel framework for family verification through facial resemblance among members. Family photo albums inevitably include photometric changes, multiple individuals, different ethnicity and genders that make it a more demanding classification problem compared to face recognition. The conventional

face recognition algorithms are designed for finding the resemblance between a pair of similar faces [36] or matching the query image with a bank of templates from an individual. More importantly, the face is always seen in the training set such as one-sample face recognition [85] or face recognition from sketches [94]. We investigate the challenging problem of unseen members to contribute to finding missing children as a critical application of family verification. We collected a database of 45 family albums from volunteers, friends and available online shared albums with total 5400 face samples. To evaluate family verification using various algorithms different scenarios were defined to cover real life applications. The proposed inter-patch, constraint-free analysis of the face segments reduces the number of required patches for family verification as tabulated in Table 3.6 and Table 3.7.

Chapter 4

A Novel Approach for Removing Redundant Feature Sets

4.1 Introduction

As aforementioned in Chapter 2 and Chapter 3, family verification not only encounters inevitable uncontrolled photometric changes but also unbalanced datasets, mixed ethnics, mixed genders, different age groups and aging effect bring up more challenges to this application. The higher intra-class variation of family members' face manifolds requires more distinctive feature operators to extract as much as information embedded in faces. Consequently, there would be fewer number of discriminant features among high dimensional extracted features by conventional operators or feature selection methods. This necessitates removing redundant features from selected features or adjusting feature parameters for extraction. Reducing feature dimension would speed up the overall family verification process. Furthermore, reducing the computational load in the training stage would benefit critical applications such as finding the missing member in surveillance systems to prevent child abduction in which the system is in urgent need of fast training.

In this chapter, we enhance feature's distinctiveness for family members' faces and adjust parameters of the proposed feature operator and selected conventional operators. The proposed method, not only enhances family verification performance compared to conventional features, but also converges faster and removes redundant feature sets that may even deteriorate the performance [66].

To answer the above shortcomings and tackle challenges of family verification, we illustrate our overall method to verify family members by incorporating redundant feature set removal in section 4.2. Then, we propose the Uniformly-sampled Thresholds for LBP (UTLBP) to extract the information embedded in family members' faces in section 4.3. Distinctiveness of the proposed operator is then compared to the state-of-the-art feature operators.

Although the information carried out by the proposed operator is enhanced but the feature dimension increases as the price we pay. We analyse the conventional redundant feature removal algorithms and discuss their shortcomings and invalid assumptions that do not satisfy family verification facts in section 4.4 and 4.5 respectively. At first, we propose the Genetic Algorithm (GA) searching algorithm with a novel chromosome structure to analyse the information carried by the feature sets in section 4.6. We also propose to remove redundant feature sets by diversity analysis in section 4.7. This chapter ends with experimental results and discussions on family verification performance improvement compared to conventional face recognition algorithm depicted in Figure 4.1. The proposed method converges 10 times faster on average by comparing the error obtained at each iteration. It also enhances family verification performance compared to the state-of-the-art features with significantly smaller feature dimension.

Figure 4.1: A general Face Recognition block diagram [6]

Overall, the proposed method reduces the enhanced feature operator dimension by removing redundant feature sets (adjustment of feature parameters) while getting the same performance of all feature sets. Finally, the proposed feature operator's distinctiveness is measured with conventional feature operators on the large CAS-PEAL database.

4.2 Our Proposed Approach for Family Verification

Nowadays, many successful algorithms for face data processing are learning-based [2]. A general block diagram of the state-of-the-art face recognition methods is depicted in Figure 4.1 [6]. As aforementioned in section 3.5, inevitable photometric changes and larger intra-class variability makes the family class boundaries highly non-linear compared to face verification. As a result, family verification requires,

- **More distinctive feature operators,** Due to less amount of resemblance among family members, common feature operators designated for face recognition may not be informative enough to extract minute similarity details of members' faces. This necessitates designing more discriminative features to extract more information embedded in family faces.

- **Removal of redundant features,** Among the pool of extracted features fewer features are required for family verification due to less intra-class similarity. Hence, we expect to have more redundant information in features selected by conventional feature selection methods.
- **Speed up training stage,** In most applications, training stage of face recognition is performed offline and the testing stage benefits from large offline computations. However, in family verification we need algorithms with fast training stage or with the ability of parallel processing. Assume the case that the missing kid's family asks the shopping center security team to look for their kid using the surveillance system. Their faces need to be trained fast to adjust parameters for testing video frames in real-time and prevent the possible child abduction.

Based on the above discussions and family members' resemblance valid assumptions, we propose a novel method with *more distinctive feature operator*. The block diagram of the proposed family verification method is shown in Figure 4.2. The main steps are:

1. Extract features from faces and divide features based on feature operator parameters denoted as "Feature Extractor #i" in Figure 4.2.
2. During training stage:
 - a. Select dominant features and classify each feature set to family or non-family
 - b. Compute feature sets diversity and remove redundant features. It yields to adjust the weight of each feature set, regardless of feature selection and classification algorithm
3. During testing stage, only remaining features are computed and used as the input to bank of classifiers.
4. The decision from the fusion algorithm determines if the query face image belongs to the family or not.

Figure 4.2: The overall framework of redundant feature set removal steps

Therefore, regardless of classifier and feature operators, the whole block diagram should be trained on the family database to remove the redundant information and recognize family members based on their common facial features. Family verification performance firstly relies on distinctive operators to extract minute similarity features from the face. As the first step, we illustrate the Uniformly-sampled Thresholds for LBP feature operator.

4.3 Enhanced Feature Operator for Family Members' Face Description

To overcome the nonlinearity of face manifolds a local appearance-based feature space, such as Haar [28], Gabor wavelet-based features [40] and Local Binary Pattern (LBP) [95] can be employed on the image [2]. In the literature, several features are proposed to extract useful information of the face for further processing such as Haar [28], Gabor [40] and Local Binary Pattern [95]. Haar features will usually produce higher feature dimension compared to Gabor. Typically, for an image of size 20×24 , we get about 70,000 Haar features but only 19,200 Gabor features with a bank of 40 jets are generated. It has been reported that Gabor features are robust to small variation in illumination and alignment [6] compared to Haar features. However, the

bank of Gabor jets and parameters need to be optimized for applications and datasets [96], [97]. Next, we review LBP feature operator modifications due to their low complexity, robustness and reported accurate performance for face data analysis [44, 45, 56, 98].

4.3.1 Local Ternary Patterns (LTP1)

There are several improvements in the literature for the LBP operator to embed more facial information to the operator. The function $s(x)$ in equation (2-9) fails to distinguish between the patterns generated from neighboring pixels with equal value versus those with larger pixel levels. The Local Ternary Pattern (LTP), is defined to address this problem while maintaining the LBP feature structure and capabilities [99]. The function $s_{ltp}(x)$ returns three distinctive results (smaller, equal and larger) when comparing the pixel values as,

$$s_{ltp}(x) = \begin{cases} 2, & x > 0 \\ 1, & x = 0 \\ 0, & x < 0 \end{cases} \quad (4-1)$$

4.3.2 Local Triplet Patterns (LTP2)

The $s_{ltp}(x)$ function is quite sensitive to noise and illumination changes and is then changed to a more general function considering an interval of ± 5 in the x conditions denoted as s_{LTP} in equation (4-2) [44]. As a result, a sequence of triplet codes (in base 3) leads to a large and sparse histogram of feature values. It reduces the error slightly than the conventional LBP by considering a scaling parameter that combines histogram bins to reduce them to 64 or 256 bins [99].

$$s_{LTP}(x) = \begin{cases} 2, & x > 5 \\ 1, & 5 > x > -5 \\ 0, & x < -5 \end{cases} \quad (4-2)$$

To tackle this problem they proposed a scaling and a neighboring parameter for the feature [44]. By using the scaling parameter, the extracted feature is scaled down and quantized to a smaller limit. However, two major drawbacks draw our attention in the proposed method,

1. Different scales led to poorer or better results in separate conditions that a constant value could not be easily generalized for all variations.
2. By using the neighboring parameter, the LTP levels calculated from a sequence of the triplet codes selected from less than ' P ' neighboring pixels reduces the LBP distinctiveness and generality.

Overall, the experiments conducted showed the superiority of LTP to LBP although in some experiments they achieved the same or poorer performance compared to LBP while the feature dimension and required memory space are increased.

4.3.3 Local Directional Patterns (LDP1)

In this research work, authors utilized neighboring pixel's gradient magnitude along a specific direction instead of the raw intensity value and then implemented the conventional LBP algorithm [98]. They computed the Kirsch edge response equal to 3 and used its output to encode the image texture by LBP. In fact, the approach taken in [98] is a pre-processing stage before conventional LBP implementation values in different directions by setting the filter parameter equal to 3 and used its output to encode the image texture by LBP.

4.3.4 Local Derivative Patterns (LDP2)

Zhang et al. explored the feasibility and usefulness of using high-order local patterns for face representation. In this scheme, LBP is considered as the non-directional first-

order local pattern operator since LBP determines all-direction first-order derivative binary results and misses more detailed information enclosed in the input object [45]. Local Derivative Patterns (LDP) was proposed to encode directional pattern features based on local derivative variations. It encodes the higher-order (n^{th}) derivative information which contains more detailed discriminative features that the first-order local pattern (LBP) fails to extract from an image as claimed by the authors [45].

4.3.5 Our Proposed Uniformly-sampled Thresholds for LBP (UTLBP)

A closer and deeper look at the LBP operator stated in equation (2-9) shows that, there are two types of information that exist at the right part of the equation but vanish at the left side of equation (2-9),

- the pixel value of the center pixel ' g_c ',
- the sign and magnitude of difference between the neighboring pixels and the center pixel, ' $(g_p - g_c)$ '.

The pixel value information ' g_c ' is sensitive to variations such as lighting or background and that is why we design features to make the classification/recognition robust to these changes. The conventional LBP operator quantizes the relative intensity information of the surrounding pixels to the center pixel. In this way, large variations of the intensity level do not change the LBP feature considerably since the relative gray scale information of neighbor pixels is almost unchanged. However, the quantization error of the conventional LBP feature is very high due to step function $s(x)$ with one level quantization at "0" level. We thus propose to include the missing information using Uniformly-sampled Thresholds for LBP (UTLBP) as explained below. Note that, the pixel depth ' Pix_D ' is equal to "255" for our analysis below although other pixel depth values are applicable.

The pure relative information of the surrounding pixel value to the center pixel can be embedded into the LBP operator by the term ' $g_p - g_c$ '. If we omit the step function to keep the pure grayscale information ' $g_p - g_c$ ' the feature evaluated for a 3×3 window would be as large as 2×256^8 or approximately 2×10^{20} , (2 is multiplied due to sign bit). To tackle this problem, we consider Uniformly-sampled Thresholds for LBP (UTLBP) operator to evaluate the binary code for each threshold. Thresholds are uniformly distributed within the range $(-Pix_D < Thr < Pix_D)$. The feature vector of pixel 'c' is modeled as,

$$UTLBP_{P,R}(c) = [utlbp_{P,R}(c, Thr_1), \dots, utlbp_{P,R}(c, Thr_T)] \quad (4-3)$$

where,

$$utlbp_{P,R}(c, Thr_k) = \sum_{p=0}^{P-1} s(g_p - g_c - Thr_k) 2^p \quad (4-4)$$

Each feature subset is extracted by employing ' Thr_k ' where $1 \leq k \leq T$. Note that, increasing number of LBP neighborhood pixels from P to Q [30], enlarges feature dimension by order of $2^{(Q-P)}$ while increasing number of thresholds from T to T_l enlarges the dimension by $(T_l - T)$ times.

4.3.6 UTLBP Operator Distinctiveness Analysis

The UTLBP's informativeness and distinctiveness is compared to the conventional LBP in Table 4.1. Assume that LBP features of "original images" shown in the 2nd column are to be extracted. The specific portion of the image is magnified in the 3rd column and the enter pixel grid mask is drawn on images to simplify LBP feature

extraction for the reader. Magnified images are shown in higher resolution to be visible and down-sampled images were used for feature extraction.

Images #1 and #2 show samples of an African woman and a Caucasian man with mustache and beard, respectively. Although to human perception they are totally different the LBP feature extracted from the lips' corner are totally the same. Images #3 and #4, show light blue and dark brown samples among variety of eye colors whose conventional LBP features are the same. Features from the grid center point of 4 samples in Table 4.1 using UTLBP are extracted in the last column with uniformly sampled thresholds ($-30 < Thr < 30$) and the step of 15. The UTLBP feature vector is represented by features of each threshold in a row. The UTLBP feature differentiates between image #1 and #2, the two totally dissimilar human faces and eye colors in image #3 and #4.

4.3.7 UTLBP Sampling Step and Redundant Feature Sets

Utilizing multiple thresholds to improve the encoding algorithm embedded in the LBP algorithm leads to larger feature dimension. However, the order of feature dimension increase is considerably smaller than increasing the radius (R) [30]. The family verification reliability of individual Gabor wavelets and features sets extracted by each individual threshold of the proposed UTLBP operator are depicted in Figure 4.3. Due to higher intra-class variability of family members' faces, there are Gabor wavelets or thresholds with very low reliability in Figure 4.3. On the contrary, very few Gabor wavelets perform less accurate than 50% for face recognition of each member. To obtain each feature set's reliability the AdaBoost classifier is implemented on each individual feature set. Those with less than 50% accuracy are depicted in yellow.

Table 4.1: Image samples for feature extraction and comparison of the extracted UTLBP and the conventional LBP

#	Original image	Magnified portion	Conventional LBP feature	Proposed UTLBP
1			01011000	[11111110 11111000 01011000 00001000 00000000]
2			01011000	[11111110 11111110 01011000 01011000 01010000]
3			11110111	[11110111 11110111 11110111 11100001 11000000]
4			11110111	[11111111 11111111 11110111 11110111 11100111]

Figure 4.3: (a) Gabor Wavelets reliabilities (b) selected UTLBP thresholds reliabilities for family verification (Test 1) of a random family dataset. Blue bars are the feature sets with reliability of higher than 50%. Yellow bars correspond to feature sets with less than 50% reliability.

For UTLBP feature sets in Figure 4.3 (b) only thresholds of the interval $[-70, 65]$ are considered out of the whole interval $[-255, 255]$. The rest of UTLBP thresholds contain very few reliable feature sets. Moreover, there are other thresholds more reliable than the original LBP (threshold=0), that vary from family to family and need an automatic algorithm to select them. Next, we give a summary of redundant feature removal algorithms in the literature and analyze their advantages and shortcomings.

4.4 Conventional Redundant Feature Removal Methods

In general, features are extracted from the face sample and passed through the feature selector. The feature selector reduces the dimension of the extracted features and the classifier classifies the face sample into genuine or imposter classes. However, for family verification, due to higher intra-class variability of family members' faces, we propose to identify and remove selected redundant features. By removing redundant features, not only will the memory size of the stored data and the computational time decrease, it could also provide better classification results as some redundant data may deteriorate the performance [66], especially since there is only limited resemblance features. To reduce dimension of selected features, several approaches are proposed in the literature. In general search strategies can be categorized into 3 general methods as [5],

Complete search methods, which consist of exhaustive and non-exhaustive strategies. An exhaustive search may evaluate all available subsets. It may also do a “breadth first” strategy, in which the searching is stopped as soon as an optimal subset is found.

Heuristic search procedures, which are subdivided into “forward selection”, “backward selection”, “combined forward and backward” and “instance-based” categories, that are subjected to stability and may not converge [57].

Random search procedures, are divided into two groups based on the probability of the subset. That probability may be kept constant throughout the searching procedure. In the other group, the subset probability is modified as the program runs.

In our application, we encounter large number of features extracted from faces and complete search methods are computationally expensive. We discussed the shortcoming and the possibility of heuristic methods instability. Hence, we could benefit Random and complete search methods if the computational load fits available computing power.

4.5 Valid Assumptions of Features' Information Analysis

We summarize assumptions and strategies in most of feature selection algorithms in the literature under which one can only get a sub-optimal solution [100] or they do not fit human family facts as below,

- **Feature information independence for feature removal**, To benefit from the information of the new feature set, during the AdaBoost learning process [55], mutual information between the candidate weak classifiers and the new selected weak classifier is examined in [11]. The process is in fact a forward selection algorithm and it requires pre-setting the threshold for mutual information comparison. In another work, FloatBoost, the recent classifier is removed if any previously added classifier contributes to error reduction less than the latest addition [27]. The result is a smaller feature set with the same classification accuracy but this forward selection search algorithm makes the computational cost 5 times more expensive which is not applicable for cases with large number of features or image size [11]. These forward and backward selection based algorithms suffer from lack of features' joint correlation information as in [11] and [27].

-
- **Feature extraction around visually salient points**, In this strategy, features are only extracted around visually salient points [12]. There is no assurance that all fiducial patches are detected correctly or the presence of unique facial features such as mole is not as dominant as other parts of the face.
 - **Minimization of total error**, in this case there is no guaranty to minimize FNR in applications with unbalanced datasets or low number of genuine samples such as family verification.
 - **Weak cost functions**, Using Gabor features extracted with different filters parameters, Gokberk et al. proposed to use Genetic Algorithm, (GA), to localize feature jets from a pool of Gabor features and test validity of fiducial point discrimination power [100]. However, their approach requires presetting parameters for GA to converge as well as the weak fitness function. Same weaknesses are observed for the approach proposed in [97]. The fitness function of each chromosome, e.g. 1-NN classifier, a weak classifier [101], must be as fast and simple as possible so that it does not become a bottleneck for the algorithm. However, with only a few genuine samples, the total error minimization almost omits the False Negative (Rejection) Ratio portion from the total error.
 - **Percentage of reliable feature sets**, on the contrary to face verification, the percentage of reliable feature sets to verify family members is low, especially for different thresholds of UTLBP as depicted in Figure 4.3.

4.6 Our Proposed Strategy for Redundant Feature Set Removal by GA

Genetic Algorithms (GA) are a family of computational models that mimic the metaphor of natural evolution. It provides better solutions as indicated in [102] compared to other searching algorithms such as Simulated Annealing. They belong to random searching method as described before. Genetic Algorithms (GA) work to find a subset of solution pools that minimizes an objective function. They do so by encoding all available subsets by a chromosome-like data structure. By applying genetic operations such as cross over and mutation, the probability of encoded subset changes to the direction that potential solutions have more weights. This process leads to evolution of populations of individuals that are better fitted to the objective.

We propose a multiple classifier system for its advantages in statistical difference between train and generalization data as well as training data set. We assume that the most ' b ' informative feature sets are to be selected among total number of ' N ' extracted feature sets. The overall approach is as follows:

1. Feature sets (e.g. obtained from different Gabor Wavelet parameters) are extracted from the cropped face data
2. AdaBoost is used to train the classifier using feature sets taken from a particular parameter.
3. From all AdaBoost classifiers trained in step 2, an ensemble of ' b ' classifiers is chosen. The ensemble is selected by the joint diversity [92] calculation.
4. Genetic Algorithm is used to provide a fast selection. A novel chromosome structure called "Real-valued chromosome structure" is proposed to eliminate problem of invalid chromosome flooding the

population. The objective function is chosen as the correlation measurement of diversity.

5. Combine outputs from selected ensemble of classifiers using the Chair and Varshney rule [7].

4.6.1 Our Proposed Chromosome Structure and Cost Function for Genetic Algorithm

To overcome analysis of the large feature dimension based on valid assumptions, we opt for the GA to find a subset of solution pools that minimizes an objective function. As aforementioned, feature independence, total error minimization and weak cost functions were weaknesses of various algorithms in the literature. We prevent the above deficiencies of the state-of-the-art algorithms in the design of our proposed approach. The GA parameters, chromosome structure, cost function, cross over and mutual operations are explained in details below,

Real-valued Chromosome Structure, Here the phenotype of each individual in the population will be a set of ‘ b ’ feature sets. In the Binary-valued chromosome structure each feature set can be presented by the binary code of its order. To prevent invalid chromosome the acceptance rate for ‘ b ’ sets is considered as,

$$R_{accept} = \left(\frac{N}{2^q}\right)^b \quad (4-5)$$

where q is calculated from $\min\{2^q > N\}$. In this case, even by using the population size of 3000 individuals, it is only possible to run up to $b = 3$ out of 40 feature sets, (the acceptance rate of chromosome is then $(0.625)^3 = 0.2441$). The other issue is, there may be two q -bit blocks in the chromosome identical or the corresponding filter is repeated twice. However, by employing the Integer-valued Chromosome Structure only up to 20 filters can be analyzed with the accessible computational power and the repeated filter index is still problematic.

We propose a novel encoding scheme called “Real-valued chromosome structure” to deal with the two aforementioned problems. A particular individual is encoded by a chromosome, which is a string of real valued numbers in the range of (0, 1). The decoding procedure is as followings,

- Assume that $Chrom = \{c_1, c_2, \dots, c_b\}$ is the chromosome encoding an ensemble of ‘ b ’ sets out of N feature sets, where $c_i \in (0, 1)$.
- Initialize a selection $Sel = \{1, \dots, N\}$ containing all possible sets that the first set in the ensemble can take.
- The allele c_1 corresponds to set f_1 , with:

$$f_1 = Sel(1 + \lfloor c_1 \times length(Sel) \rfloor) \quad (4-6)$$

where $\lfloor x \rfloor$ indicates the largest integer number smaller than ‘ x ’. In equation (4-6), f_1 is considered as the set having the order $1 + \lfloor c_1 \times length(Sel) \rfloor$ in the selection set denoted by ‘ Sel ’.

- Update the selection set as $Sel = Sel - \{f_i\}$, the updated selection set contains all values that the following allele can be encoded.
- Repeat initialization steps for all other alleles: $c_2, c_3, c_4, \dots, c_k$.

We use a population of size 3000 and evolve it for 10 rounds as the compromise between the accuracy and computational cost. The minimum correlation value of 10 rounds is then selected.

Objective Functions, We choose the ensemble correlation measure as the objective function since we are looking for highly diverse ensembles.

Crossover Operations, the discrete recombination is applied for its simplicity. It creates the offspring by randomly choosing values from their parents for each allele, with the same probability.

Mutation Operations, for mutation of a real-valued population, the breeder genetic algorithm [103] is employed.

Employment of parallel Genetic Algorithm, due to huge number of ensembles having very low correlation, especially for those with high number of sets, a population of 3000 members cannot cover all of them. Consequently, sub-optimal selection occurs. To tackle this problem, 10 different threads of Genetic Algorithm are appointed. Each thread evolves a population of 3000 individuals over 10 generations. After 10 generations, we take 10% of the population (300 individuals) and combine them together. Thus, we have a total of maximum 3000 ensembles in the combined set. Since, the ensembles taken from different Genetic Algorithm threads may be identical; we may have less than 3000 ensembles. From these 3000 ensembles, we apply the direct search to select the ensemble with the highest accuracy.

Finally the performance of each set of classifiers is evaluated by the FNR and FPR. The final decision of the set of classifiers is made by the Chair-Varshney rule to find out if the sample belongs to the family. The pseudo code of the Chair-Varshney rule implemented on each set of classifiers is shown in details in Figure 4.4. The first algebraic term $\log(P_1/P_0)$, the threshold, is calculated using the desired FPR_D .

4.7 Our Novel Approach for Redundant Feature Set Removal

We proposed to obtain more informative facial patches in Chapter 3 for family resemblance extraction among members. We suggest applying the same method to

remove redundant feature sets. However, the number of feature sets is more than considered facial patches and computational complexity of the diversity search increases in order of power of 2.

Decision making on GA selected feature sets

- Calculate the FNR and FPR of each classifier based on the training data
- Fix the maximum desired FPR, shown as FPR_D
- Using the FPR_D , find fusion threshold from equation (3-11)

$$f = \sum_{s_i} \log \frac{1 - FNR_i}{FPR_i} + \sum_{s_0} \log \frac{FNR_i}{1 - FPR_i}$$

- Calculate the ensemble FNR from equation (2-13) and (2-14)
- Repeat steps above for all ensembles in the ensemble set selected by Genetic Algorithm
- Choose the ensemble having the lowest FNR.

Figure 4.4: The Pseudo code of the Chair-Varshney rule implementation on GA selected feature sets' corresponding classifiers

As aforementioned, diversity of a set of classifiers must be minimum for the best combination of classifiers to reach the best performance. Therefore, its inverse is maximum. The inverse of the diversity can be written as,

$$\rho_b^{-1} = 1 - \frac{1}{b} + \frac{N_T - N^t}{bN^f} \quad (4-7)$$

which is proportional to,

$$\rho_b^{-1} = 1 - \frac{1}{b} + \frac{N_T - N^t}{bN^f} \propto \frac{N_T - N^t}{N^f} = \frac{N_T - N^{tf} - N^{tn}}{N^{ff} + N^{fn}} \quad (4-8)$$

where we split N^f to N^{ff} , given by number of cases that all classifiers are false when the data is a family member. Similarly, N^{fn} is given by number of cases that all classifiers are false when the data is a non-family member. Similarly for N^t where N^{tf} is the number of cases that all classifiers are true when the data is a family member and N^{tn} , number of cases that all classifiers are true when the data is a non-family member. In our experiments, approximately all feature sets are highly reliable on non-family data with very few error samples. Therefore N^{fn} may vary between 0 to cte. (a small value), and $N_T - N^{tn}$ is the total number of family samples due to high reliability of classifiers for non-family dataset. So it can be simplified as equation,

$$\frac{N_T - N^{tf} - N^{tn}}{N^{ff} + N^{fn}} \cong \frac{N_f - N^{tf}}{N^{ff} + \text{const.}} \quad 0 \leq \text{const.} \leq 10 \quad (4-9)$$

dividing by N_f gives:

$$\frac{1 - P^{tf}(C_1 \cap C_2 \cap \dots \cap C_N)}{P^{ff}(C_1 \cap C_2 \cap \dots \cap C_N) + \text{const.}} \quad 0 \leq \text{const.} \leq \frac{10}{N_f} \quad (4-10)$$

where U_i is the set of i_{th} classifier output, P^{tf} is the probability of verifying family members correctly by all classifiers, P^{ff} is the probability of verifying family members incorrectly by all classifiers. Based on the reliability of classifiers (1-FNR), assuming ($N=2$) the range of intersection of incorrectly classified family members of two classifiers denoted by ' $\cap_{ff}$ ' is

$$0 < \cap_{ff} < \min(N_A^{ff}, N_B^{ff}) \quad (4-11)$$

where the subscript of N^{ff} is the index of classifiers. The error probability (Pe) is then evaluated by $Pe = Pe(A) + Pe(B) - Pe(A \cap B)$. The intersection is optimized by finding the most diverse classifiers and the summation of remaining positive terms $Pe(A)$ and $Pe(B)$ necessitates classifier A and B to be selected of highly reliable classifiers. This can be extended to N classifiers with the error probability of,

$$P_e = \sum_{r=1}^N P_{U_r} + \sum_{s=2}^N \left\{ (-1)^{s-1} \sum_{t=1}^{C(N,s)} P_{\{B_t\}} \right\} \quad (4-12)$$

where B :{all permutations of intersections of s classifiers out of N }, $C(N,s)$ is the permutation of s out of N , ' P ' is the probability of error and ' P_e ' is the probability of fusion error. Thus, it is required to specify which feature set corresponding classifier is reliable and what the threshold of reliability is. An appropriate threshold for the classifier reliability, ' Thr_R ', is found when the set of ' Thr_R ' reliable feature sets reaches the diversity value of all features diversity and also get the same performance as they do. The complexity of finding the threshold is just to find the most ' m ' maximum reliable feature sets. If the set of selected set satisfies the criterion then choose the minimum reliability in the selected feature set. We can easily ignore other less reliable sets by making their weights to zero. To avoid data outlier, those classifiers with " ϵ " less than Thr_R are considered, too. " ϵ " is a small value, adjusted based on the closeness of reliable filters, here is equal to "0.02". Once the group of reliable sets, N_m , is obtained, the least diverse classifiers contain the most informative feature sets for the particular family.

4.8 Experimental Results and Discussions

Our focus is to enhance *feature operator distinctiveness* and *remove selected feature redundancy* to achieve faster algorithm while getting higher performance for family verification. To compare feature's discrimination power, we utilized AdaBoost as the feature selector. We compare Haar [28], Gabor [40], conventional LBP [95], LTP [44], LDP [98] and the proposed UTLBP with two sets of thresholds. In the first set thresholds are selected uniformly between -255 and 255 with equal steps of 10 denoted as UTLBP-10 and in the second set with steps of 5 denoted as UTLBP-5. Hence the size of UTLBP features per image increases by 102 and 51 times compared

to LBP for UTLBP(5) and UTLBP(10) features that results in 40392 and 20196 features, respectively.

The Haar features are computed using the first 4 templates as in [28] which generates approximately 70,000 features. To extract Gabor features, the bank of 40 Gabor wavelets with 5 scales and 8 orientations are convolved with the image and the magnitude of the complex value is represented as the Gabor feature. Each LBP based operator is calculated with 4 different settings: 1) histogram of extracted LBP features denoted as ‘H’, 2) the resulting histogram mapped to uniform patterns denoted as ‘H-U2’, 3) feature of all pixels are stored as a vector of size (1×396) denoted as ‘Pi’ and 4) the LBP features of all pixels mapped to uniform patterns is stored as a vector of size (1×396) denoted as ‘Pi-U2’. The LBP window is limited to a 3×3 neighbor or $(R,P)=(1,8)$.

4.8.1 Performance Comparison Using Selected Features

In the first experiment, we compare discrimination power of various feature operators for family verification using the conventional face recognition method shown in Figure 4.1. For every dataset, extracted features from the image are fed to an AdaBoost classifier and the error is averaged for all families. Figure 4.5 shows the FNR, and total error (Err) of Haar [28], Gabor [40], LBP [95], LTP [44], LDP [98] and the proposed UTLBP feature with two different threshold adjustments, UTLBP-10 and UTLBP-5, for test 1, 2a and 2b after 200 iterations.

The results show that UTLBP features extracted from all pixels (Pi), outperform LBP [95], LDP [98], LTP [44] operators, Haar [28] and Gabor [40] features by 3.2%, 2.8%, 4%, 1% and 3.4% respectively. In test 1, since all members are present in the training dataset, the FNR is less than the other test scenarios which is expected.

Table 4.2: Vector dimension of extracted features from images of size 20×24 (pixel depth = 256)

Index	Feature Type	Dimension	Index	Feature type	Dimension
#1	LBP (H-U2)	59	#9	UTLBP-5 (H-U2)	6018
#2	LBP (H)	256	#10	UTLBP-5 (H)	26112
#3	LBP (Pi-U2)	396	#11	UTLBP-5 (Pi-U2)	40392
#4	LBP (Pi)	396	#12	UTLBP-10 (Pi)	40392
#5	UTLBP-10 (H-U2)	3009	#13	Gabor	19200
#6	UTLBP-10 (H)	13056	#14	Haar	74000
#7	UTLBP-10 (Pi-U2)	20196	#15	LTP (Pi)	396
#8	UTLBP-10 (Pi)	20196	#16	LDP (Pi)	396

The features dimensions are tabulated in Table 4.2 for image size of 20×24 and pixel depth equal to 256. Although the feature size is larger than LBP [95], LTP [44] and LDP [98], it is still comparable to Gabor [40] and is lower than Haar [28] features.

4.8.2 Redundant Feature Set Removal Comparison

In this section, the performance of the proposed methods to remove redundant feature sets is compared with the conventional algorithms in terms of computational time and accuracy. Since the proposed scheme consists of two main blocks, classifier fusion and redundant feature removal, we first analyze the performance of the proposed individual feature set classifier fusion using optimal decision making rule with the conventional face recognition algorithm shown in Figure 4.1 on both Gabor wavelets and UTLBP-10 features. Among four settings of UTLBP-10, (H-U2, H, Pi-U2 and Pi) features of all pixels without mapping (Pi) are selected for further experiments and analysis.

Figure 4.5: The average family verification total error (Err) and False Negative Ratio (FNR) of selected features for Test 1, Test 2.a and Test 2.b using the conventional face recognition method shown in Figure 4.1

Table 4.3: Family verification average FNR and FPR comparison using the conventional face recognition algorithm shown in Figure 4.1 and the proposed individual feature set classifier fusion

Feature type	Test	Error type	Conventional Face			Proposed Method		
			Iterations			Iterations		
			5	10	100	5	10	100
40 Gabor Wavelets	T1	FNR	0.70	0.685	0.56	0.55	0.56	0.52
		FPR	0.031	0.033	0.039	0.031	0.033	0.039
	T2.a	FNR	0.84	0.83	0.8	0.77	0.75	0.75
		FPR	0.02	0.02	0.022	0.02	0.02	0.022
	T2.b	FNR	0.84	0.83	0.814	0.78	0.77	0.75
		FPR	0.022	0.023	0.026	0.022	0.023	0.026
UTLBP-10	T1	FNR	0.521	0.499	0.393	0.429	0.43	0.383
		FPR	0.034	0.033	0.024	0.034	0.033	0.024
	T2.a	FNR	0.69	0.69	0.65	0.62	0.65	0.615
		FPR	0.027	0.022	0.023	0.029	0.022	0.023
	T2.b	FNR	0.75	0.75	0.72	0.70	0.70	0.66
		FPR	0.031	0.03	0.022	0.031	0.03	0.022

Table 4.3 shows the family verification results obtained with these two features. To compare the rate of convergence, we compute the FNR and FPR at the 5th, 10th and 100th iterations. The results show that the proposed scheme is able to achieve the performance of conventional face recognition algorithm shown in Figure 4.1 at 100th iteration after just 5 to 10 iterations.

The average computational time of the total training and testing stages of 40 Gabor Wavelets, UTLBP-10 and UTLBP-5 features using both Genetic Algorithm and the proposed individual feature set classifier fusion is tabulated in Table 4.4. The machine uses the Intel Q9400 2.67 GHz CPUs. It shows that the computational time is comparable for the same number of iterations. The results of computational time comparison in TABLE III shows that,

- The classifier fusion does not increase the computational cost
- Since the proposed method requires less number of iterations to achieve comparable performance, the proposed approach incurs lower computational time than the face recognition algorithm shown in Figure 4.1.
- Moreover, it consists of fully computational independent batches of classifiers that can run on parallel machines to speed up the training stage.

The overall computational time of running on parallel machines comprises of dividing the whole computational into multiple tasks and the time required to run each task on a worker machine. Creating tasks for all worker machines on the server with Intel Q9400 2.67 GHz CPU takes around 8 seconds depending on feature operator. Hence, the total computational time required to train multiple classifiers is equal to total computational time in Table 4.4 divided by available worker machines plus 8 seconds of task division. Note that, maximum number of parallel jobs (worker machines) is limited to number of feature sets. Finally, we run another series of experiments to find the achievable classification accuracy if the number of features used is reduced to ‘ b ’ most informative feature sets for the Gabor wavelets and UTLBP features.

Table 4.4: The average computational time comparison of the proposed method and conventional face recognition in Figure 4.1 in different iterations.

Classifier type	Iterations	Computational time (sec)			Parallel Computing
		Gabor	MTLBP-10	MTLBP-5	
Conventional Face Recognition	5	48	53	71	Not applicable
	10	97	104.5	138	
	100	1006	1052	1400	
Optimum Decision maker	5	42	48	67	Runs on parallel machines
	10	93	101	134	
	100	997	1041	1366	

We compared the proposed feature removal scheme with the popular Principal Component Analysis (PCA) approach and the proposed Genetic Algorithm (GA). We compute the total error of test 1, after 10 iterations, for $b=2, 3, \dots, 18$ out of N feature sets and the results are plotted in Figure 4.6.

The experimental results obtained show that using the proposed method UTLBP-5 and UTLBP-10 we achieve the accuracy of all feature sets at $b=5$ and $b=8$ respectively while it happens at $b=18$ for Gabor features. As from $b=18$ the performance remains the same for all features and algorithms selected in Figure 4.6 the curves are plotted up to 18 feature sets. Employing the Genetic algorithm as the method for feature removal to achieve all feature sets performance occurs at $b=27$. Based on the results, number of features can be significantly reduced using the proposed redundant feature removal scheme without significantly affecting the performance when all features are used. In addition, the proposed scheme outperforms both PCA and Genetic Algorithm in terms of accuracy achievable with equal number of feature sets.

Figure 4.6: The total error for the set of b most informative sets for different features using the proposed two-stage redundant feature removal, PCA and GA

4.8.3 UTLBP Performance on Face Recognition

We also examine the performance of the proposed UTLBP operator on face recognition in controlled environment. Among face recognition databases [2], CAS-PEAL provides large-scale face images with different sources of variations [13]. It contains 30,900 images of 1,040 subjects. Among them, 379 subjects have images with 6 different expressions. 438 subjects have images wearing 6 different accessories. 233 subjects have images between 10 to 31 lighting changes. 297 subjects have images against 2 to 4 different background colors. Furthermore, 66 subjects have images recorded in two sessions at a 6-month interval to provide aging effect. It also includes 21 poses for each subject. To prepare samples, faces are aligned and cropped to size 80×95 using provided eyes positions. Samples taken under controlled environment of PEAL database are grouped in Figure 4.8 for illustration.

Figure 4.7: A facial image divided into 5×5 regions (b) the weights set for the weighted χ^2 dissimilarity measure. Black squares indicate weight 0.0, dark gray 1.0, light gray 2.0 and white 4.0.

As proposed by Ahonen et al. in [30], faces are divided into regions (we selected 25 regions) as shown in Figure 4.7 (a). LBP histograms of regions are extracted. Using the proposed method in [104] and the weight (w_i) for each region is calculated as shown in Figure 4.7 (b). The weighted Chi square χ^2 dissimilarity measure is employed as the distance measurement of the nearest neighbor classifier as in equation (4-13) [104],

$$\chi_w^2(x, \xi) = \sum_{j,i} w_j \frac{(x_{i,j} - \xi_{i,j})^2}{x_{i,j} + \xi_{i,j}} \quad (4-13)$$

where ‘ x ’ and ‘ ξ ’ are the normalized histograms to be compared and indices ‘ i ’ and ‘ j ’ refer to i -th bin in histogram corresponding to j -th local region ‘ j ’ with the weight ‘ w_j ’ calculated for that region. The CAS-Peal database contains Gallery (1040 images) and Training sets (1200 images) which are used as training samples [13]. The Training set contains four images randomly selected from the frontal subset of 300 random subjects. Each sample in the probe set is tested against all training samples to find the nearest neighbor. If the individual is recognized incorrectly it is considered as an error.

We choose aging, lighting, background, accessory and distance probe sets of the CAS-Peal database as our work does not include pose correction and expression recognition. Among illumination normalization methods suggested in [13] as the standard procedure, histogram equalization is selected for image preprocessing.

The histogram of 25 regions are calculated using LBP(1,8), LBP(2,12), LTP(1,8), 2nd order Local Derivative Patterns (LDP) and the proposed feature operator UTLBP(1,8) and UTLBP(2,12). To adjust thresholds, 200 random individuals of the FERET dataset [17] were tested for face recognition. The UTLBP features extracted by each threshold in the interval $(-254 < Thr < 254)$ with the step equal to 5 were employed separately to perform face recognition for each threshold reliability analysis. The range of highly reliable thresholds were chosen that are $(-69 < Thr < 69)$ with equal step of “5” for UTLBP. The weighted Chi square χ^2 distance of UTLBP is measured for each threshold histogram subset and summed together. The recognition error rates of feature operators for selected probe sets are tabulated in Table 4.5.

Figure 4.8: Some example images in CAS-PEAL-R1 Database adapted from [13]

Table 4.5: Recognition error rate for feature operators in each probe set of the CAS-PEAL database

Probe sets	Feature types recognition error					
	LBP(1,8)	LBP(2,12)	LTP(1,8)	LDP (2 nd order)	UTLBP(1,8)	UTLBP(2,12)
Aging	0.37	0.30	0.59	0.37	0.34	0.22
Lighting	0.90	0.886	0.926	0.888	0.865	0.873
Background	0.034	0.012	0.157	0.032	0.014	0.009
Accessory	0.06	0.032	0.216	0.06	0.032	0.022
Distance	0.051	0.047	0.18	0.054	0.021	0.011

We selected the first 1000 samples and the first 500 of the lighting and accessory probe sets respectively. All samples of other given probe sets were selected for evaluation. As shown in Table 4.5, UTLBP(2,12) outperforms other selected LBP features in the literature. Compared to LBP(2,12), the significant amount of improvement (8%) occurs in the aging probe set which is one of the common changes in family. These variations such as accessory, lighting, distance and background almost exist in family photo albums in which UTLBP achieves higher accuracy.

We also tested the performance improvement of UTLBP to state-of-the-art descriptors on texture classification. The UTLBP-10 achieves 94% and 61.6% accuracy on Outex 12 [105] and KTH-TIPS2b [106] respectively. The accuracy of conventional LBP is 87% and 57.8 on Outex 12 [105] and KTH-TIPS2b [106] respectively.

4.8.4 UTLBP Features Discrimination Power on Family Data

The same algorithm as in [30], the weighted Chi square χ^2 distance for the nearest neighbor classification is utilized to determine distinctiveness of the proposed

UTLBP features on family datasets. The family verification error is evaluated in Table 4.6 for selected features for test 1 and 2. The results show that UTLBP(2,12) reduces the total error (Err) by almost 1% compared to LBP(2,12) in test 1. The 2nd order LDP was reported to outperform LBP(1,8) which is mostly consistent with Table 4.5 and Table 4.6. The interesting point is, we expect to achieve higher accuracy by increasing the LBP operator radius ' R ' and consequently number of sampling pixels ' P '. In test 2.a and 2.b, where we encounter a more challenging situation when a member is missing enlarging the radius from "1" to "2" leads to higher total error (Err) by 2% in test 2.a. Among *selected features in the literature*, LBP(2,12) achieved the least error rate in face recognition as stated in Table 4.5.

Table 4.6: Family verification error evaluation for selected feature operators

Feature type	Error types for Test 1			Error types for Test 2.a			Error types for Test 2.b		
	Err	FPR	FNR	Err	FPR	FNR	Err	FPR	FNR
LDP (2 nd order)	0.0500	0.0321	0.2067	0.1159	0.0255	0.5775	0.0900	0.0251	0.6587
LBP (1,8)	0.0518	0.0406	0.1470	0.1020	0.0317	0.4482	0.0906	0.0276	0.6223
LBP (2,12)	0.0483	0.0403	0.1281	0.1224	0.0422	0.4954	0.1029	0.0387	0.6568
UTLBP (1,8)	0.0506	0.0338	0.1471	0.0925	0.0171	0.4470	0.0802	0.0141	0.6214
UTLBP (2,12)	0.0391	0.0316	0.1312	0.1015	0.0246	0.5038	0.0910	0.0216	0.6616

Compared to LBP(2,12), the proposed feature UTLBP(1,8) could reduce the family verification total error (Err) in case of the missing member by 3% and 2% in test 2.a and 2.b, respectively. Note that the same thresholds of UTLBP in Table 4.5 were employed for feature extraction of family datasets.

4.9 Concluding Remarks

This chapter explains our proposed method to enhance feature operator's distinctiveness for family face description and remove redundant feature sets to reduce the computational load. As the first step, we propose the UTLBP feature operator by reducing the quantization error of the step function. Consequently, the feature dimension increases as the order of thresholds employed in the feature. Next, we analyze the shortcomings and invalid assumptions of the conventional redundant feature selection algorithms. Based on the facts of family members and valid assumptions, we propose two different approaches to remove redundant feature sets to reduce computational time. Firstly, we define a novel chromosome structure to search for "b" number of the most informative feature sets. The second approach searches among feature sets combinations and determines the least diverse sets that achieve the highest accuracy. As a result, the overall proposed method enhances family verification performance compared to the state-of-the-art operators with significant feature dimension reduction as stated in Figure 4.5 and Figure 4.6. We also examine the proposed UTLBP feature distinctiveness on the CAS-PEAL database that outperforms other similar operators in the literature as the results tabulated in Table 4.5.

Chapter 5

A New Framework for Utilizing Similarity of Family Member's Image Segments

5.1 Introduction

In this chapter, we study valid assumptions and analyze psychological studies on facial patch resemblance among family members. Facial resemblance could be measured easier when larger face samples are available and the computational power could handle required computations. We propose a novel method to estimate each member's facial patch similarity and consolidate the available information of facial similarity to recognize family members more accurately, compared to the state-of-the-art fusion methods. Our proposed approach is also able to benefit from human brain understanding from facial similarity among members for family verification.

In this chapter, we assume that large face samples are available to separate patches from the whole face with enough resolution so that human brain could distinguish the amount of resemblance in each individual patch. The facial resemblance is the

key solution to recognize family members, particularly if enough information from each individual patch is available. However, in case of large face samples, processing the similarity among family members varies in three directions compared to the state-of-the-art face processing algorithms:

1. Facial similarity depends on gender and individuals in the family, as diverse results on the resemblance of daughters, sons and their corresponding fathers or mothers are reported [107]
2. Distinctive resemblance features and the amount of similarity differs in facial segments of a particular family [108]
3. The kinship verification is not processed using overall spatial information in the human brain. The accuracy achieved by combining “masked” facial patches is higher than considering the full face [35, 78]

As a result, the algorithm must exploit the *resemblance* of each family *member's* image *segment* to perform family verification as illustrated in Figure 5.1. The key contributions of this chapter are divided into two major directions,

1. Estimate the amount of resemblance among family members that differs from member to member and image segments
2. Combine the estimated family members’ facial resemblance based on the given member’s image segment information to classify the query image

From left to right at Figure 5.1 (a) the members’ average face of a family dataset and their overall resemblance evaluated in chapter 3 are depicted, respectively. In this chapter, we estimate each member’s facial segment resemblance to the whole family (shown in Figure 5.1 (b)) based on their segmented faces depicted in Figure 5.1 (c). In section 5.2, we study recent psychological studies on the resemblance among family members and the performance of human brain in recognizing family

members. As a result, the amount of facial resemblance of family members depends on each member's facial patch. The possibility of analysis and embedding such information was missing in the state-of-the-art fusion methods. We summarize our proposed approach to benefit from each member's facial patch resemblance information for family verification in section 5.3. The state-of-the-art score fusion methods are studied in section 5.4 to choose the best type for family verification based on the valid assumptions and the cost function to be minimized in family verification. Our proposed approach is explained in details in section 5.5 to consolidate the available information of facial patch resemblance of each member. We also propose to estimate the amount of similarity of each individual member's facial patch resemblance to obtain required information for fusion task. The proposed method is then tested in different scenarios against state-of-the-art score fusion rules that achieves considerable improvement to recognize family members. At the end of section 5.6, we illustrate the online survey conducted to obtain human brain understanding from facial resemblance of each member's facial patch.

Figure 5.1: (a) The average face of family members and their overall facial patch resemblance (b) each member's facial patch resemblance to be considered in the family verification algorithm (c) family members' facial patches split using the golden ratio mask template implemented in general

The experimental results also show that our proposed approach could estimate the similarity slightly more accurate than human perception. This chapter ends with the extension of term ‘family’ to families of objects. We examine our proposed approach for score fusion with the facts of family resemblance on the family of objects available in the Caltech-256 database. We also tested the proposed fusion rule on the family of two objects, motorcycle and car side-view in the available Caltech-256 database that achieved up to 9% accuracy improvement.

5.2 Family Members' Facial Parts Similarity in Psychological Studies

Psychologists asked participants to assess facial pictures of nonfamily samples individuals to guess the resemblance between children and parents [107] and also among siblings by Maloney and Dal Martello [35] or adult faces [109]. It was discovered that other people's judgments about facial resemblance referred as “social mirror” affects individuals behaviors as in spouse/child abuse [82] or paternal investment [80]. There are interesting consensuses from psychological studies on facial resemblance for family verification as follows,

- **It depends on gender and members.** There is differential resemblance between the two parents, depending on children's gender that suggests the facial phenotype might be towards one parent as a reaction to costs and benefits of paternal investment [107]. Alvergne et al. deduced from the collected dataset that the facial resemblance inverts for boys, but not for girls [107]. However, the effect of children's age and gender for preferential resemblance needs to be investigated carefully in a large sample size, assuming it can be generalized.

-
- **No general rule can be extended to all families and members.** Christenfeld and Hill implied that the infants' resemblance is more towards their father [110] although it was then explored by Bre'dart and French who claimed that their experiment was "A failure to replicate Christenfeld and Hill (1995)" [111]. Moreover, studies of resemblance in children only [112], or the similarity of interest of two members such as father and children [111] or three members (mother, father, and children [113] did not come to a common decision of precise similarity measurement among members.

 - **What is the rank of facial segments containing family resemblance information?** the question was firstly raised by Maloney & Dal Martello as "Where are kin recognition signals in the human face" utilizing children's facial segments without considering their genders [35]. They inferred that the lower half of children's faces including mouth and chin shape carries less useful information about genetic family resemblance, probably due to growth through childhood. Debruine et al. investigated finding facial similarities of family members in adult faces in [109] as the continuation of [35].

 - **What are the mechanisms and algorithms for assessment of family facial patches similarities?** It was addressed in research work that algorithms and feature types of kinship recognition signals are poorly known [107, 109]. Based on studies, recognizing faces may result from facial organs arrangement processing ability [114]. Research work on family facial resemblance [78] reveal that the resemblance extraction is not processed using overall spatial information such as ratio of the distance between face organs as in [12]. In reality, the performance of family verification improved when face regions were analyzed instead of the "full-face" as examined by Dal Martello and Maloney [35]. They concluded "combination of kinship information from the

two halves of the face can be treated as *optimal combination of independent cues*” [35].

Similarly, when humans observe family members, we usually hear e.g. “this family is all blond”, “all siblings have connected eyebrows like their father” or “even the boys have got sharp noses like their mom and sister”. These sayings also suggest that resemblance in a family differs in facial parts and varies between family members. In another words, it is member and patch specific. As a result, the family verification strategy should consider facial resemblance of each individual family member’s image segment (facial patch) to the whole family to satisfy the assumptions and important findings from the above discussion.

5.3 Our Proposed Framework

Our goal is to enhance the family verification accuracy based on family facial parts resemblance assumptions discussed in Section 5.2. As aforementioned, the dependence of family resemblance *to gender and image segments* and the method to *benefit from such a priori information* were missing in previous studies. Hence, our proposed approach is to,

- estimate individual member’s image segments similarity
- combine the estimated resemblance based on the a priori information

We utilize the a priori information of “member” and “the corresponding image segment” in decision making which play important roles for verification of belonging to the family. We consider the face mask developed in Chapter 3 from the original facial beauty analysis [37] to split the face regions with slight changes as shown in Figure 5.2.(c). The main steps of our proposed algorithm are as follows:

1. Split the face to regions based on the developed facial mask as in Figure 5.2.(c)
2. Extract features from the face patches using the most discriminate operator among selected features
3. Process features extracted from each patch to get the match score
4. During the training stage:
 - a. Compute every present member's patch similarity to the whole family
 - b. Adjust the threshold to perform fusion of selected patches' match scores by incorporating every member's patch similarity measurement
5. During the testing stage, match scores of each patch are computed and used as the input to the proposed fusion rule
6. The decision from the combined output determines if the face image belongs to the family or not.

We explain the feature operator and selection algorithms in section 5.6 in which experiments are illustrated in details. Our main contribution relies on the estimation of individual member's facial segment similarity and to benefit from this information to improve family verification performance. Next, we analyse shortcomings of the conventional methods and provide valid assumptions as the basis of our proposed method.

Figure 5.2: Aligned and Cropped family faces are split into regions (b) the golden ratio mask template indexing implemented in Chapter 3 (c) modified patches selected in this chapter

5.4 Conventional Score Fusion Rules

Among fusion methods at different stages of recognition, score fusion is usually preferred due to the balance between information possession and the fusion complexity [89]. We need to investigate which method suits the task of family verification. We analyze different score fusion methods and discuss their shortcomings and advantages as follows,

- **Classifier-based approaches**, the output scores are concatenated to form a final vector and a second stage of classification is performed to obtain the final results of being an imposter or a genuine sample. It is complicated to include minimization of error types in the classification based algorithms [48]. Hence, in case of finding missing member which is very critical to fail to spot a genuine sample (*minimize P_{miss}*) classifier-based approaches are not suitable.
- **Transformation-based approaches**, match scores are converted to a normalized common domain which requires large amount of data for evaluation [115]. Though, in the application of family verification, we encounter unbalanced and small number of training datasets for some members due to limited sources of genuine family data samples.
- **Density-based approaches**, these score fusion methods employ likelihood ratio criterion. It guarantees optimal performance if the match score densities are known and well estimated. The genuine and imposter match score densities were estimated in [89, 116] for score fusion. One of the main advantages of density based approaches is the minimum False Negative Ratio (FNR) achieved at the desired False Positive Ratio (FPR) that suits family verification. An extensive test on the large NIST database [117] concluded that the density-based method could achieve the highest accuracy among

selected algorithms [117]. The estimation complexity of the kernel density approach was then reduced by modeling the PDF as finite Gaussian Mixture Model (GMM) of joint PDFs in [89]. Their proposed method slightly outperformed the Support Vector Machine (SVM) classifier on the output scores [89].

We choose the density-based classifier fusion approach as it achieves the minimum (FNR) at the desired (FPR) that suits family verification while outperforming other existing score fusion methods.

5.5 Consolidation of Family Member's Facial Parts Resemblance

Family verification is a two class problem. We need to formulate the family verification problem to derive the fusion rule based on the human family members' facial parts similarities assumptions. The task is to verify if the query individual (Y) with the given sample X_Y belongs to p -th family denoted by (F_p) or $(Y \in F_p?)$. Each dataset has parents and at least one child (not infant). Each family dataset is labeled and annotated to M_p members where n_{pq} is the number of total images for q -th member from p -th family. The face sample of q -th member from p -th family denoted as X_{pq} is divided into L regions. Features of each region are extracted and given to AdaBoost matcher to calculate the match score. The vector of all patches' match scores is $S = [S_1, \dots, S_L]$. Let $f_g(S)$ be the conditional joint density of the L match scores given the p -th family genuine samples, and $f_i(S)$ be the conditional joint density of the L match scores given the imposter samples. Now let us consider this problem from the Bayesian perspective. According to Bayes decision theory, we should decide $Y \in F_i$ if

$$P(Y \in F_i | X_Y) > P(Y \notin F_i | X_Y) \quad (5-1)$$

or,

$$\frac{P(Y \in F_i | X_Y)}{P(Y \notin F_i | X_Y)} > 1 \quad (5-2)$$

According to Bayes theorem,

$$P(Y \in F_i | X_Y) = \frac{P(X_Y | Y \in F_i) \cdot P(Y \in F_i)}{P(X_Y)} \quad (5-3)$$

and

$$P(Y \notin F_i | X_Y) = \frac{P(X_Y | Y \notin F_i) \cdot P(Y \notin F_i)}{P(X_Y)} \quad (5-4)$$

By substituting equations (5-3) and (5-4) in equation (5-2), the decision rule can be simplified as, decide $Y \in F_i$ if

$$\frac{P(X_Y | Y \in F_i) \cdot P(Y \in F_i)}{P(X_Y | Y \notin F_i) \cdot P(Y \notin F_i)} > 1 \quad (5-5)$$

We assume equal prior probabilities, (i.e. $P(Y \in F_i) = P(Y \notin F_i) = \frac{1}{2}$) then the decision rule depends only on the likelihood ratio, which is defined as

$$LR_i(X_Y) = \frac{P(X_Y | Y \in F_i)}{P(X_Y | Y \notin F_i)} \quad (5-6)$$

Based on the Neyman-Pearson theorem, T(S) is the optimal test such that,

$$T(S) = 1 \quad \text{if} \quad \frac{f_{g,p}(S)}{f_i(S)} \geq \eta \quad (5-7)$$

where, vector ' S ' is the match score of the query face ' Y '. We define the False Positive Ratio, (FPR) as

$$P(T(S) = 1 | H_0) = FPR \quad (5-8)$$

The Neyman-Pearson (N-P) guaranties that, given the desired FPR, the optimal test for deciding whether the match score vector S is a genuine user or an impostor is given by equation (5-7). We adjust the threshold η for a fixed FPR such that the likelihood ratio test minimizes the False Negative Ratio, FNR . Here, we need to estimate the joint density functions $f_{g,p}(S)$ and $f_i(S)$. Due to complexity and sensitive parameter adjustment of KDE [117], we select the Gaussian Mixture Model (GMM) estimation [10, 89]. Hence, the N-P test $T(S)$ is performed on the GMM estimated PDFs as,

$$T(S) = 1 \quad \text{if} \quad \frac{\hat{f}_{g,p}(S)}{\hat{f}_i(S)} \geq \eta \quad (5-9)$$

Dal Martello and Maloney observed that the kinship information obtained from the two halves of the face can be considered as "optimal combination of independent cues". Among selected patches patch 1 and 2 are usually altered by hair particularly in ladies faces that reduces the correlation between them, significantly. The only similar patches are patch 3 and 4 that include eyes. However, since the dependence characteristics among genuine and imposter left and right eyes do not differ, the independence assumption of eyes patches' match scores does not alter the fusion performance [10]. Nandakumar et al. showed that in this condition, embedding the dependence information to the fusion scheme will not lead to significant performance improvement [10]. They also indicated, as estimation of each marginal density is more accurate and efficient than estimation of the complex, multi-dimensional joint density function of the combined random variables, the independence assumption would ease the match score fusion, especially when we encounter *lack of genuine data*

as in family verification. Hence, we extend equation (5-9) to the product of marginal density functions,

$$\frac{\hat{f}_{g,p}(S)}{\hat{f}_i(S)} = \frac{\prod_{r=1}^L \hat{f}_{g,p,r}(s_r)}{\prod_{r=1}^L \hat{f}_{i,r}(s_r)} = \prod_{r=1}^L \frac{\hat{f}_{g,p,r}(s_r)}{\hat{f}_{i,r}(s_r)} \quad (5-10)$$

The numerator of equation (5-10) estimates the marginal density functions of the genuine samples regardless whose patch scores are given but we know that the family resemblance of each patch differs from member to member. To deploy members' patches information and dissimilar resemblance among members' facial patches we propose to extend the nominator of equation (5-10) based on the *Total probability* and Bayes formula as,

$$\frac{\hat{f}_{g,p}(S)}{\hat{f}_i(S)} = \frac{\prod_{r=1}^L \hat{f}_{g,p,r}(s_r)}{\prod_{r=1}^L \hat{f}_{i,r}(s_r)} = \prod_{r=1}^L \frac{\sum_{q=1}^{Mp} \hat{f}_{g,p,r,q}(s_r | Y \in U_q) * P_g(Y \in U_q)}{\hat{f}_{i,r}(s_r)} \quad (5-11)$$

where U_q is the set of all samples belong to q -th member and ' Y ' is the given face. We know that $\hat{f}_g(S)$ and $\hat{f}_i(S)$ are the short forms of conditional density functions $\hat{f}_p(S | Y \in F_p)$ and $\hat{f}(S | Y \notin F_p)$ where F_p is p -th family. Hence, the estimated genuine density function is equal to,

$$\begin{aligned} \hat{f}_p(S | Y \in F_p) &= \prod_{r=1}^L \hat{f}_{p,r}(s_r | Y \in F_p) \\ &= \prod_{r=1}^L \sum_{q=1}^{Mp} \hat{f}_{p,r,q}(s_r | \{Y \in U_q\}, \{Y \in F_p\}) * P(Y \in U_q | \{Y \in F_p\}) \end{aligned} \quad (5-12)$$

The amount of similarity between q -th member's r -th patch and the whole family is now deployed in equation (5-12). Next, we illustrate how to estimate the amount of similarity for each member's facial patch. Hence, the final N-P test is equal to,

$$\begin{aligned}
\frac{\hat{f}_p(S | Y \in F_p)}{\hat{f}(S | Y \notin F_i)} &= \frac{\prod_{r=1}^L \hat{f}_{p,r}(s_r | Y \in F_p)}{\prod_{r=1}^L \hat{f}_r(s_r | Y \notin F_p)} \\
&= \frac{\prod_{r=1}^L \sum_{q=1}^{M_p} \hat{f}_{p,r,q}(s_r | \{Y \in U_q\}, \{Y \in F_p\}) * P(Y \in U_q | \{Y \in F_p\})}{\prod_{r=1}^L \hat{f}_r(s_r | Y \notin F_p)}
\end{aligned} \tag{5-13}$$

5.5.1 Estimation of Similarity

We propose to estimate each member's patch similarity by computing the probability of q -th member's r -th patch samples to belong to family P. Given all training q -th member's r -th patch samples, the GMM marginal density function of the genuine $\hat{f}_{p,r,q}(s_r | \{Y \in U_q\}, \{Y \in F_p\})$ and imposter $\hat{f}_{i,r}(s_r)$ samples are estimated. A Bayesian classifier is investigated to calculate the probability of recognizing the given sample to belong to family p . The parameters of the Bayesian classifier and estimated $P(Y \in U_q | \{Y \in F_p\})$ are stored for the test stage. The overall proposed method to utilize family resemblance based on valid assumptions discussed in section 2 is summarized in the following pseudo code in Figure 5.3.

5.6 Experimental Results and Discussions

Firstly, we need to determine the distinctive feature operators for facial patches and then implement the fusion rule on the rendered match scores for family verification.

5.6.1 Split Faces into Patches

There are various types of information in the human face that could contribute to the family verification. Genetic similarities in the family include skin and hair color, face shape, face organs and minute features of the face.

Consolidation of facial resemblance

Split the given face into L patches

Extract features of the patches

Obtain the match scores from each patch $S = [S_1, \dots, S_L]$

$A=0; M=1;$

for patch $r=1:L$

for member $q=1:M_p$

 read the similarity measurement $P(Y \in U_q | \{Y \in F_p\})$ calculated

 compute $sum = \hat{f}_{p,r,q}(s_r | \{Y \in U_q\}, \{Y \in F_p\}) * P(Y \in U_q | \{Y \in F_p\})$

$A = sum + A$

End for

$Ratio = A / \hat{f}_{i,r}(s_r)$

$M = Ratio * M$

End for

Compare M with the threshold obtained in training

Figure 5.3: Pseudo code of the proposed method to utilize individual member's facial patch similarity

We opt for the developed facial mask in [108], designed based on the face golden ratio to select regions containing the genetic similarities of the face. However, there are patches as shown in Figure 5.2 (b), which do not carry much face texture. Moreover, the informativeness analysis in Chapter 3 shows that the information carried by these patches was previously captured by other patches. Hence, we do not consider patches 1, 4, 6, 8, 12 and merge patch 9, 10 and 11 to form the nose region. Patch 16 and 17 are also merged to form the chin region to get the final mask as shown in Figure 5.2 (c).

5.6.2 Feature Extraction

To compare feature's discrimination power on each patch, we utilized AdaBoost as the matching algorithm and set the threshold to default value '0'. We compare Gabor [6], LBP [30] and DSIFT [54] operators, due to their robustness against illumination changes, misalignment and acceptable feature size for large face real time processing [40]. To extract Gabor features, the bank of 40 Gabor wavelets with 5 scales and 8 orientations are convolved with the image and the magnitude of the complex value is represented as Gabor feature. Due to image size, 16 sample points around center pixel with radius equal to 2 are selected for LBP to extract LBP(16,2) features [30]. DSIFT features were extracted using the finest settings for feature extraction based on the image size with block size equal to 10 pixels. Features of each patch are separated for each facial region shown in Figure 5.2 (c). The Total Error (Err) and FNR of selected features for every patch in Test 1, 2a and 2b are shown in Figure 5.4, after 25 iterations of AdaBoost classifier training. The results show that Gabor features outperform other selected features in most patches in terms of total error and FNR.

5.6.3 Score Fusion Rule

In this section, the performance of the proposed method to fuse the match scores obtained from AdaBoost matchers is compared with the state-of-the art GMM fusion rule [10]. We use the provided MATLAB code provided in [118] to estimate the density functions. The algorithm automatically estimates number of components and their parameters [118]. For every family, the training data match scores rendered by AdaBoost are utilized to estimate their density functions. The Bayesian classifier is trained for each member's patch to adjust parameters for the testing stage.

Figure 5.4: The average family verification total error (Err) and False Negative Ratio (FNR) of selected features for (a) Test 1, (b) Test 2.a and (c) Test 2.b on facial patches

The Receiver operating characteristic, ROC, curve is the most representative curve to compare fusion algorithms performance. The ROC curves of the proposed fusion rule show the achieved improvements in Figure 5.5 (a), (b) and (c) for Test 1, 2.a and 2.b, respectively. The ROC curve is plotted as Genuine Positive Ratio (GPR) versus FPR. The proposed method outperforms the state of the art score fusion rule up to 17%, 12% and 10% in Test 1, 2.a and 2.b, respectively.

The family verification performance of the proposed fusion rule is also compared with the SVM classifier. The (Radial Basis Function) RBF kernel SVM parameters need precise adjustment which cannot be accurately calculated on the available training dataset based on the criteria given in [90]. We employed the linear SVM on the match scores from facial patches. Hence, the SVM classifier is trained on the same training dataset and evaluated using the same testing set using the code provided in [90]. The results obtained at equal FPR are tabulated in Table 5.1. We select the minimum average FPR achieved by the SVM classifier to compare the GPR achieved by the algorithms. As we expected, algorithms employing kernel functions may not perform well in case of unseen data [2]; as in Test 2.a and 2.b SVM GPR is less than the proposed method, particularly it is the least accurate in Test 2.b.

Table 5.1: Performance comparison of the SVM classifier, the proposed method and the algorithm suggested in [10]

Scenarios		Test 1		Test 2.a		Test 2.b	
Error types		FPR	GPR	FPR	GPR	FPR	GPR
Method	Proposed method	0.042	77%	0.04	67%	0.036	57%
	Linear SVM	0.042	76%	0.04	60%	0.036	44%
	Nand. et al. [10]	0.042	73%	0.04	57%	0.036	54%

(a)

(b)

(c)

Figure 5.5: The ROC curves for the proposed fusion approach and the fusion rule by Nandakumar et al. [10] for (a) Test 1, (b) Test 2.b (c) Test 2.b

5.6.4 Utilizing Human Perception

Psychologists conduct various experiments to discover the human ability and artificial intelligent researchers use it as a reference to compare computer understanding with human brain. Feng et al. ask 16 participants to compare their proposed kinship method with human perception [12]. We asked 30 engineers and researchers from 4 different regions East Asia, Middle East, Europe and US to rate the similarities for 10 selected families. A sample survey of the first family from selected faces of George Bush family is illustrated in Figure 5.6.

(a)

	A lot	Very	So so	A bit	No Ideal
Forehead *	☐	☐	☐	☐	☐
Eyes *	☐	☐	☐	☐	☐
Nose *	☐	☐	☐	☐	☐
Cheek *	☐	☐	☐	☐	☐
Mouth *	☐	☐	☐	☐	☐
Chin *	☐	☐	☐	☐	☐

(b)

Figure 5.6: (a) samples of Bush’s family dataset samples for the designed survey
(b) the survey questions and multiple choices answers samples

Participants were asked to estimate how much family members resemble each other in their forehead, eyes, nose (including cheeks), mouth and chin. We designed the online survey to analyze the human perception of each member's patch resemblance to the whole family to perform fusion rule by incorporating human brain understanding of member's patch similarity estimation. Particularly, in test 2.a and 2.b we encounter a critical situation and we need to deploy all available information to find the missing member.

However, the conducted survey allows participants to guess accurately the resemblance of the missing member's patches, since there are samples of the missing member in the provided album, while the algorithm just uses the information of available members. The choices of similarity amount were provided as "a lot", "very", "so so", "a bit" and "No idea" in the survey and mapped to vector $[4, 3, 2, 1, 0]$ respectively. The average of all participants' choices mapping is calculated for the member's patch similarity measurement for each family in the survey.

We substitute the human perception in equation (5-11) to compare the performance of family verification results in Test 1, 2.a and 2.b. ROC curves of the proposed method fusion rule, human perception of family resemblance and the score fusion rule proposed in [10] are plotted in Figure 5.7 (a), Figure 5.7 (b) and Figure 5.7 (c), respectively. The estimated member's patch similarity slightly outperforms the human perception on 10 selected family datasets except in Test 2.a that their performance is very close.

We include selected samples of the provided genuine samples of 3 different scenarios with selected corresponding test samples for further illustration in Table 5.2. The output of the proposed method is compared with the approach suggested by Nandakumar et al. [10].

Figure 5.7: The ROC curves for the proposed fusion approach using human perception, the proposed resemblance estimation method and the fusion rule by Nandakumar et al. [10] for (a) Test 1, (b) Test 2.b (c) Test 2.b

Table 5.2: Outputs of the proposed method and the proposed method and the algorithm suggested in [10] on selected samples of 3 family albums for the 3 different scenarios

Provided genuine data
Test type	Test 1		Test 2.a		Test 2.b	
Genuine query samples
True class	Family	Family	Family	Family	Family	Family
Results of the proposed method	Family	Family	Family	Family	Family	Imposter
Results of Nandakumar et al. [10]	Imposter	Family	Family	Imposter	Imposter	Imposter

5.7 Families of Objects

We also verify the enhancement achieved in the proposed match score fusion rule, by considering family of objects from the publically available dataset of objects, Caltech-256 [119]. The object family members need to have at least 3 members to represent a family of various and diverse members but with image patch similarities. Moreover, each member must have enough samples for training of the patch resemblance estimation. As we demonstrated patches' match scores fusion, solid objects datasets with minor change of view point were chosen from Caltech-256. Among objects in

Caltech-256, we selected “Motorcycles” and “Car side-view” sets as two different families of object. Samples with complex background were removed to reduce parameters that may affect the accuracy. In order to have conclusive results 10-fold cross validation was performed to obtain the average accuracy. Two experiments were conducted,

- Motorcycles were divided into Touring, Racing, Old, Jumping and Normal (low cc) members with 70, 101, 66, 31 and 56 samples respectively. The approximate positions of the wheels centers were extracted [120] and images were aligned and cropped to size 100×170 as shown in Table 5.3. K-fold cross validation was performed by generating 10 families randomly with 30 samples from each member.
- Car images captured from side view were divided into 2-door coupe, Van & pick up and 4-door sedan cars with 16, 18 and 71 samples respectively. The approximate positions of the wheels centers were extracted [120] and images were cropped to size 65×190 as shown in Table 5.3. K-fold cross validation was carried out by generating 10 families randomly with 15, 15 and 35 samples from 2-door coupe, Van and 4-door sedan members, respectively.

Non-family members were chosen from all mountain bicycles, touring bikes and clutter samples comprising 1000 samples. Non-family samples were resized based on the motorbike or car tests. There are minute misalignments in samples because either wheels are usually occluded by various body parts or slight changes of view makes circles more like oval. We choose DSIFT features as SIFT descriptors are robust against change of view point and slight displacement [54]. The step size of DSIFT features is adjusted to 11 for detailed feature extraction. The human facial parts locations are mostly consistent to other individual and that makes it easier to divide the face into precise regions that include particular facial patch.

Table 5.3: Samples of each motorcycles and car side-view family members

Object	Member	Samples
Motorcycle	Touring
	Racing
	Old
	Jumping
	Normal (low cc)
Car side-view	Two-door coupe
	Van & Pick-up
	4-door sedan

However, due to various models and diverse designs of motorbikes and cars, we divide the object image into three horizontal and two vertical equal sections to obtain six equal patches. The proposed match score fusion algorithm is employed for the six regions scores. The ratio of training and testing sets is considered '1/2'. The ROC curves of Test 1 on 10-fold motorcycle and car side-view recognition are plotted in Figure 5.8 (a) and Figure 5.8 (b), respectively.

The proposed fusion rule could improve the recognition accuracy by 9% and 3.5% in low FPR for car side-view and motorcycle recognition, respectively compared to [10]. We believe that the larger difference of improvement obtained in the car family is due to more diverse members than motorcycles. We conducted the object recognition experiments on the Caltech-256 database due to its availability. However, the performance improvement will increase if more variety of members such as, scooter and tricycle in the motorcycle family or mini and old cars in the car side-view family are provided.

5.8 Concluding Remarks

This chapter explains our proposed method to incorporate the amount of resemblance of each individual member's patch to perform family verification. Researchers in the field of psychology conducted various tests to find the resemblance among family members. Our analysis and important consensuses of psychologists reveal that the facial resemblance among family members differs from member to member and is facial patch specific. As the first step, we propose to fuse the amount of resemblance of each individual member's patch to the whole problem of family verification using segmented facial patches. As the score level fusion is preferred in the literature due to its wealth of information and low data dimension, we embed the available a priori information in the score fusion rule.

(a)

(b)

Figure 5.8: The ROC curves of match scores fusion rule comparison for the proposed approach and Nandakumar et al. [10] using Caltech-256 (a) motorcycle and (b) car side-view datasets

Next, we propose to estimate the amount of resemblance for each patch, based on the prior information of the member. The experimental results show that the proposed method outperforms the state-of-the-art score fusion rules by considering valid assumptions on family members' facial patch similarity. We also conduct an online survey to obtain human perception of the facial resemblance between family members. Our proposed method of facial resemblance estimation slightly outperforms human perception. Moreover, we extend the term 'family' to objects. The results obtained in Figure 5.8 shows that we could enhance object recognition by employing the amount of resemblance of object subsets 'family members' on selected families of objects from CALTECH-256 database up to 9%.

Chapter 6

Conclusion and Recommendations

In this thesis, we have presented multiple proposed frameworks that seamlessly extend the previous works of face data processing to be more accurate and robust and less computationally expensive for family verification. We also presented experimental results demonstrating the effectiveness of our proposed framework for recognizing family members, which, to the best of our knowledge, has not yet been studied in detail in the computer vision literature. We explained each of our frameworks as a unified analytical approach to solve the challenges associated with family verification. In this final chapter, we summarize the contents of this thesis and outline its contributions. We follow this summary with a discussion of various open research problems that come to light as a result of this thesis.

6.1 Conclusion

In the first part of this thesis, we introduce our framework for family verification through an analysis of facial resemblance among members. To the best of our

knowledge, the general problem of recognizing members of the same family has not yet been analyzed and studied in the field of computer vision. The main applications of family verification in our daily life include recognizing missing family members by means of resemblances found in available face data and the benefits of commercial products that arise from recognizing family members. Currently, family photo albums are the only available sources of family face data for performance evaluation. They inevitably include photometric changes, multiple individuals, and different ethnicities and genders, which make it a more demanding classification problem compared to face recognition. On the other hand, face recognition algorithms are designed to match the query image with a bank of templates from an individual. More importantly, the face is always seen in the training set. We also investigated the challenging problem of unseen members in order to contribute to finding missing children as a critical application of family verification. We created a database of 5400 face samples by collecting 45 family albums from volunteers, friends, and shared albums available online. In order to evaluate family verification using various algorithms, different scenarios were defined covering real life applications. In the first step, we extract the facial resemblance among family faces, regardless of the feature and the matching algorithm. As family verification benefits from Mixture of Experts (MOE) due to limited genuine data and “divide and conquer method,” we propose an inter-patch, constraint-free analysis of face segments in order to reduce the number of patches required for family verification in the collected dataset.

Our second contribution involves developing a novel method for enhancing the feature operator’s distinctiveness for family face description and removing redundant feature sets in order to reduce the computational complexity. In this regard, we propose the UTLBP feature operator by reducing the quantization error of the step function. The feature dimension increases by an order of ‘T’ thresholds employed in the feature as a result, which necessitates removing excess thresholds that only increase the computational load without contributing to the performance. Next, we analyze the shortcomings and invalid assumptions of the conventional redundant

feature selection algorithms. We propose two different approaches in order to remove redundant feature sets and reduce computational time by considering the dependence of the feature sets. Firstly, we define a novel chromosome structure for the Genetic Algorithm (GA) in order to search for the minimum number of the most informative feature sets. The second approach identifies the least diverse sets that achieve the highest accuracy equal to employing all of the feature sets. We are thus able to perform family verification with less feature dimensions and obtain higher accuracy by means of the proposed UTLBP feature of higher discrimination power. We also conducted the standard tests on the CAS-PEAL database with controlled changes in terms of lighting, background, aging, and accessories. The proposed feature, UTLBP, outperforms other similar operators in the literature in identifying the face under each selected conditions.

In the third part of our contributions, the proposal of a novel method of employing the information on the amount of patch similarity of each individual member for family verification purposes. We reviewed psychological studies and analyzed the facts associated with facial resemblance among family members in order to derive a fusion model of the overall resemblance among members. We concluded that family facial resemblance differs from member to member and is facial patch specific. Firstly, we derive a score fusion model based on the amount of resemblance of the facial patch of each individual member for family verification using segmented facial patches. We incorporate the available a priori information of member's facial patch similarity into the score fusion rule. Next, we estimate the amount of resemblance for each patch using the training set in order to provide the parameters required in the testing stage. We conducted experiments on the family database created under three different scenarios. The proposed method outperforms the state of the art score fusion rules by employing the patch similarity of each family member to the overall score fusion rule. We also conducted an online survey in order to gain insight into the human brain's perception from the amount of facial resemblance between family members in each segment. Our proposed method for facial resemblance estimation

slightly outperforms that of human perception. In the end, we extended the term “family” to objects. The results of object recognition experiments for selected families of objects in the CALTECH-256 demonstrate that we can enhance object recognition by employing the amount of each segment’s resemblance of object subsets, or, in other words, of “family members.”

In the experimental works, we conclude our work by demonstrating the utility of each proposed framework by applying them to an uncontrolled database of family members containing inevitable changes in terms of photometrics, aging, accessories, and expressions. The absence of a comprehensive database of family albums in the literature implies that we have defined a new challenging problem for computer vision with critical applications to daily life.

6.2 Recommendations

This thesis gives rise to a number of interesting and promising research avenues worth pursuing. We describe several of these research avenues in the following sections of the thesis.

6.2.1 Performance Improvement of Face Recognition

The manifold or distribution of a particular class of faces is comprised of variations in the appearance of that face class appearance, whereas the manifold of imposters’ faces includes everything else. Face recognition can be considered as a task that involves discriminating between the manifolds of genuine faces and imposters’ faces. Similarly, family verification distinguishes common manifolds of faces in the overall face manifold. Figure 6.1 (a) illustrates faces manifolds belonging to individual A

versus those belonging to individual B and (b) manifolds of faces of members of a family within the entire face manifold.

We proposed multiple frameworks in order to tackle the challenges of family verification. We are able to classify the query image into family members and imposters in a photo. Various research papers imply that family members are often recognized as the query person in face recognition applications [31]. Therefore, it would be beneficial to include family members (similar faces) as family faces in the search for faces domain and to then look for the query person. As shown in Figure 6.1(b), it is difficult to distinguish between multiple persons when they look like one another in a family, particularly if they are identical twins. Hence, using family verification, we expand the search field to the red area of family members shown in Figure 6.1(b) in order to reduce the face recognition error of recognizing similar family members and in order to convert them to positive genuine family samples.

Figure 6.1: Simplified 2-Dimensional manifolds comparison for (a) face recognition adapted from [2] (b) the i^{th} face manifold in a family separated from others by the red boundary

6.2.2 Recognition of Individual Family Members

In this thesis, we proposed the more discriminative feature operator UTLBP, the redundant feature set removal algorithm, and fused the available information of the facial patch similarity of each member in order to improve family verification performance. However, as mentioned in section 6.2.1, the enhancement of face recognition relies on accurately distinguishing the individual from other family members. The overall error in state-of-the-art faces recognition comprises,

$$\text{Total Error of face recognition} = E_F + E_N \quad (6-1)$$

where,

E_F = Error of recognizing family members

E_N = Error of recognizing nonfamily members

Utilizing family verification will eliminate E_F ; the next step is to focus on features that make the query individual distinctive from other members in order to minimize E_N . The redundant feature removal algorithm identifies features that either carry the information previously captured or that do not contribute to family verification. Our second recommendation for future research is to cultivate benefit from features selected or eliminated for family verification and select dominant features to foster recognition of members of a family. Hence, features and the classifier should be designed in such a way that they are able to differentiate small differences among the huge amount of resemblance between the family members. This problem is more sensible in the context of the identical twin identification problem. We naively compared the performance of Face Recognition (FR) and Family Verification before Face Recognition (FVFR) in Table 6.1 using the approach proposed in [6]. We selected discriminate features of family verification and dominant features among family members by extracting Gabor Features and AdaBoost as the feature selector. The FNR was reduced by including dominant features of family verification for face recognition.

Table 6.1: Performance comparison of Face Recognition (FR) [6] and our recommendation to perform Family Verification prior to Face Recognition (FVFR)

Scenario	Test 2a			Test 2b		
Error of the Algorithm	FNR	FPR	Err	FNR	FPR	Err
FR	0.38	0.008	0.021	0.46	0.017	0.007
FVFR	0.30	0.006	0.016	0.40	0.017	0.006

Hence, by studying features selected or removed in family verification, we investigated which of these could contribute to recognize individuals among similar faces, or could cause error in face recognition.

6.2.3 General Object Recognition Enhancement Using Members' Image Segments' Similarities

In Chapter 5, we explain our proposed method to incorporate the amount of resemblance of each individual member's patch to perform family verification. Firstly, we estimate the amount of resemblance of each individual member's patch with regards to the whole family, and then fuse the estimated amount of resemblance to recognize the human/object family. The experimental results demonstrate that the proposed method outperforms the state-of-the-art score fusion rules by considering the valid assumptions of family members' facial patch similarity. We extend the term "family" to objects, and the results obtained in Figure 5.8 show that we were able to enhance object recognition. The proposed approach requires three major processing steps that are normal stages for face data processing, but are only applicable to certain objects.

-
1. Objects should be normalized in terms of base points that are common in the object family.
 2. Objects should be segmented to split almost identical segments, which is only applicable to rigid objects.
 3. It is only applicable to families of objects from available object databases that include multiple, distinctive enough members, such as car side-view and motorbike families and their corresponding members.

Hence, our third recommendation is to come up with a general method for improving object recognition using each member's segment similarity to the whole family for non-rigid objects. This is feasible through other feature operators' similarity analyses and extension of available objects databases in terms of sufficient number of samples per object family member.

Author's Publications

Journal Papers:

- [1] M. Ghahramani, W. Y. Yau, and E. K. Teoh, "A Novel Approach for Recognizing People of the Same Family," *International Journal of Biometrics*, December 2012, in press.
- [2] M. Ghahramani, W. Y. Yau, and E. K. Teoh, "Exploring Individual Family Member's Image Segments Similarity to Favor Family Verification," Submitted to *Machine Vision and Applications*, December 2012.

Conference Papers:

- [3] M. Ghahramani, W. Y. Yau, and E. K. Teoh, "Enhancing Local Binary Patterns Distinctiveness for Face Representation," *presented in the IEEE International Symposium on Multimedia (ISM2011)*, Dana Point, California, USA, December 5-7, 2011, pp. 440-445.
- [4] M. Ghahramani, W. Y. Yau, and E. K. Teoh, "Family facial patch resemblance extraction," *at the Proceedings of the 10th Asian Conference on Computer vision - Part II*, Queenstown, New Zealand, 2011, pp.388-399.

-
- [5] M. Ghahramani, N. M. Dang, E. K. Teoh, and W. Y. Yau, "A novel approach to remove redundant Gabor wavelets for family classification," in *The 11th IEEE International Conference on Control, Automation, Robotics and Vision*, (ICARCV), Singapore, 2010, pp. 2537-2542.
- [6] F. Saei Manesh, M. Ghahramani, and Y. P. Tan, "Facial part displacement effect on template-based gender and ethnicity classification," in *the 11th IEEE International Conference on Control Automation Robotics and Vision (ICARCV)*, Singapore, 2010, pp. 1644-1649.
- [7] M. Ghahramani, H. L. Wang, W. Y. Yau, and E. K. Teoh, "Unseen family member classification using mixture of experts," in *the 5th IEEE Conference on Industrial Electronics and Applications (ICIEA), 2010*, Taiwan, 2010, pp. 336-339.
- [8] M. Ghahramani, H. L. Wang, W. Y. Yau, and E. K. Teoh, "Feature study and analysis for unseen family classification," in *the proceedings of SPIE, the 2nd International Conference on Digital Image Processing (ICDIP 2010)*, Singapore, 2010, pp. 75460Q-6.

Bibliography

- [1] *Introduction to Support Vector Machine (SVM) Models*. Available: <http://www.dtrek.com/order.htm>
- [2] S. Li and A. Jain, "Handbook of Face Recognition," ed: Springer, 2005.
- [3] M. Aly, "Face Recognition using SIFT Features. ", Caltech, Pasadena, California, USA, Technical Report, 2006.
- [4] R. L. Hsu, "Face Detection and Modeling for Recognition," Ph.D., Michigan State University, 2002.
- [5] H. L. Manoranjan Dash, "Feature Selection for Classification," *Intell. Data Anal.*, vol. 1, pp. 131-156, 1997.
- [6] M. Zhou and H. Wei, "Face Verification Using GaborWavelets and AdaBoost," presented at the Proceedings of the 18th International Conference on Pattern Recognition - Volume 01, 2006.
- [7] Z. Chair and P. K. Varshney, "Optimal Data Fusion in Multiple Sensor Detection Systems," *Aerospace and Electronic Systems, IEEE Transactions on*, vol. AES-22, pp. 98-101, 1986.
- [8] Y. Freund and R. E. Schapire, "A decision-theoretic generalization of on-line learning and an application to boosting," *J. Comput. Syst. Sci.*, vol. 55, pp. 119-139, 1997.
- [9] G. Roberts. (27 Mar 2006) Twin sister - One black, One white - turn heads in England. *Jet*. 64.
- [10] K. Nandakumar, A. Ross, and A. K. Jain, "Biometric fusion: does modeling correlation really matter?," presented at the Proceedings of the 3rd IEEE international conference on Biometrics: Theory, applications and systems, Washington, DC, USA, 2009.

-
- [11] U. Bagci and B. Li, "Parallel AdaBoost algorithm for Gabor wavelet selection in face recognition," in *Image Processing, 2008. ICIP 2008. 15th IEEE International Conference on*, 2008, pp. 1640-1643.
 - [12] F. Ruogu, K. D. Tang, N. Snavely, and C. Tsuhan, "Towards computational models of kinship verification," in *Image Processing (ICIP), 2010 17th IEEE International Conference on*, 2010, pp. 1577-1580.
 - [13] G. Wen, *et al.*, "The CAS-PEAL Large-Scale Chinese Face Database and Baseline Evaluations," *Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on*, vol. 38, pp. 149-161, 2008.
 - [14] Z. Lichun, C. Junwei, L. Yue, and P. Wang, "Face Recognition Using Scale Invariant Feature Transform and Support Vector Machine," in *Young Computer Scientists, 2008. ICYCS 2008. The 9th International Conference for*, 2008, pp. 1766-1770.
 - [15] L. Thomas. (2009). *Jackson family anger as Oliver! star Lester claims: I'm the daddy*. Available: <http://www.dailymail.co.uk/tvshowbiz/article-1205288/Is-Mark-Lester-father-Michael-Jacksons-daughter-Paris.html#>
 - [16] A. C. Gallagher and C. Tsuhan, "Clothing cosegmentation for recognizing people," in *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*, 2008, pp. 1-8.
 - [17] R. Hietmeyer, "Biometric identification promises fast and secure processing of airline passengers," *The International Civil Aviation Organization Journal*, vol. 55, pp. 10-11, 2000.
 - [18] W. Zhao, R. Chellappa, P. J. Phillips, and A. Rosenfeld, "Face recognition: A literature survey," *ACM Comput. Surv.*, vol. 35, pp. 399-458, 2003.
 - [19] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in *Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on*, 2001, pp. I-511-I-518 vol.1.
 - [20] X. Geng, Z.-H. Zhou, and K. Smith-Miles, "Automatic Age Estimation Based on Facial Aging Patterns," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, pp. 2234-2240, 2007.
 - [21] E. Makinen and R. Raisamo, "Evaluation of Gender Classification Methods with Automatically Detected and Aligned Faces," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 30, pp. 541-547, 2008.
-

-
- [22] M. Michaels. (2010). *Alleged Michael Jackson Daughter Denied DNA Test By Court*
 - [23] A. J. Sedlak, D. Finkelhor, H. Hammer, and D. J. Schultz., "National Estimates of Missing Children: An Overview," Office of Juvenile Justice and Delinquency Prevention, Office of Justice Programs, Washington, DC: October 2002.
 - [24] T. Kanade, "Picture Processing System by Computer Complex and Recognition of Human Faces," doctoral dissertation, Kyoto University.
 - [25] S. Milborrow and F. Nicolls, "Locating Facial Features with an Extended Active Shape Model," presented at the Proceedings of the 10th European Conference on Computer Vision: Part IV, Marseille, France, 2008.
 - [26] A. Pentland, B. Moghaddam, and T. Starner, "View-based and modular eigenspaces for face recognition," in *Computer Vision and Pattern Recognition, 1994. Proceedings CVPR '94., 1994 IEEE Computer Society Conference on*, 1994, pp. 84-91.
 - [27] S. Z. Li and Z. Zhang, "FloatBoost Learning and Statistical Face Detection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 26, pp. 1112-1123, 2004.
 - [28] P. Minh-Tri and C. Tat-Jen, "Fast training and selection of Haar features using statistics in boosting-based face detection," in *Computer Vision, 2007. ICCV 2007. IEEE 11th International Conference on*, 2007, pp. 1-7.
 - [29] D.-H. Liu, K.-M. Lam, and L.-S. Shen, "Optimal sampling of Gabor features for face recognition," *Pattern Recogn. Lett.*, vol. 25, pp. 267-276, 2004.
 - [30] T. Ahonen, A. Hadid, and M. Pietikainen, "Face Description with Local Binary Patterns: Application to Face Recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, pp. 2037-2041, 2006.
 - [31] M. Das and A. C. Loui, "Automatic face-based image grouping for albuming," in *Systems, Man and Cybernetics, 2003. IEEE International Conference on*, 2003, pp. 3726-3731 vol.4.
 - [32] Z. Stone, T. Zickler, and T. Darrell, "Autotagging Facebook: Social network context improves photo annotation," in *Computer Vision and Pattern Recognition Workshops, 2008. CVPRW '08. IEEE Computer Society Conference on*, 2008, pp. 1-8.
 - [33] P. Singla, H. Kautz, L. Jiebo, and A. Gallagher, "Discovery of social relationships in consumer photo collections using Markov Logic," in *Computer*

- Vision and Pattern Recognition Workshops, 2008. CVPRW '08. IEEE Computer Society Conference on*, 2008, pp. 1-7.
- [34] D. B. K. L. M. D. P. Barclay, "A cue of kinship promotes cooperation for the public good," *Evolution & Human Behavior* vol. 29, pp. 49-55, 2008.
- [35] M. F. Dal Martello and L. T. Maloney, "Where are kin recognition signals in the human face?," *Journal of Vision*, vol. 6, 2006.
- [36] R. B. Torun, M. Yurdakul, and P. Duygulu, "Finding similar faces," in *Signal Processing and Communications Applications Conference, 2009. SIU 2009. IEEE 17th*, 2009, pp. 960-963.
- [37] S. R. Marquardt. *Marquardt Beauty Analysis*. Available: www.beautyanalysis.com
- [38] M. Turk and A. Pentland, "Eigenfaces for recognition," *J. Cognitive Neuroscience*, vol. 3, pp. 71-86, 1991.
- [39] L. Xi, H. Weiming, Z. Zhongfei, and W. Hanzi, "Heat Kernel Based Local Binary Pattern for Face Representation," *Signal Processing Letters, IEEE*, vol. 17, pp. 308-311, 2010.
- [40] C. Liu, "Gabor-based kernel PCA with fractional power polynomial models for face recognition," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 26, pp. 572-581, 2004.
- [41] T. Ojala, M. Pietikainen, and T. Maenpaa, "Multiresolution Gray-Scale and Rotation Invariant Texture Classification with Local Binary Patterns," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, pp. 971-987, 2002.
- [42] D. G. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints," *Int. J. Comput. Vision*, vol. 60, pp. 91-110, 2004.
- [43] H. Bay, A. Ess, T. Tuytelaars, and L. V. Gool, "Speeded-Up Robust Features (SURF)," *Comput. Vis. Image Underst.*, vol. 110, pp. 346-359, 2008.
- [44] T. Xiaoyang and B. Triggs, "Enhanced Local Texture Feature Sets for Face Recognition Under Difficult Lighting Conditions," *Image Processing, IEEE Transactions on*, vol. 19, pp. 1635-1650, 2010.
- [45] B. Zhang, Y. Gao, S. Zhao, and J. Liu, "Local derivative pattern versus local binary pattern: face recognition with high-order local pattern descriptor," *Trans. Img. Proc.*, vol. 19, pp. 533-544, 2010.
- [46] R. Polikar, "Ensemble based systems in decision making," *Circuits and Systems Magazine, IEEE*, vol. 6, pp. 21-45, 2006.

-
- [47] A. A. Ross, K. Nandakumar, and A. K. Jain, "Biometrics: When Identity Matters," in *Handbook of Multibiometrics*, ed: Springer, 2006.
 - [48] A. A. Ross, K. Nandakumar, and A. K. Jain, "Levels of Fusion in Biometrics," in *Handbook of Multibiometrics*, ed: Springer, 2006.
 - [49] B. Scassellati, "Eye finding via face detection for a foveated, active vision system," presented at the Proceedings of the fifteenth national/tenth conference on Artificial intelligence/Innovative applications of artificial intelligence, Madison, Wisconsin, United States, 1998.
 - [50] P. I. Wilson and J. Fernandez, "Facial feature detection using Haar classifiers," *J. Comput. Small Coll.*, vol. 21, pp. 127-133, 2006.
 - [51] Y. Adini, Y. Moses, and S. Ullman, "Face recognition: the problem of compensating for changes in illumination direction," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 19, pp. 721-732, 1997.
 - [52] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, "Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, pp. 711-720, 1997.
 - [53] B. Scholkopf, A. Smola, and K.-R. Muller, "Nonlinear component analysis as a kernel eigenvalue problem," *Neural Comput.*, vol. 10, pp. 1299-1319, 1998.
 - [54] W. Jian-Gang, L. Jun, Y. Wei-Yun, and E. Sung, "Boosting dense SIFT descriptors and shape contexts of face images for gender recognition," in *Computer Vision and Pattern Recognition Workshops (CVPRW), 2010 IEEE Computer Society Conference on*, 2010, pp. 96-102.
 - [55] Y. Freund and R. E. Schapire, "A decision-theoretic generalization of on-line learning and an application to boosting," presented at the Proceedings of the Second European Conference on Computational Learning Theory, 1995.
 - [56] T. Ahonen and M. Pietikäinen, "Pixelwise Local Binary Pattern Models of Faces Using Kernel Density Estimation," presented at the Proceedings of the Third International Conference on Advances in Biometrics, Alghero, Italy, 2009.
 - [57] J. Bang-Jensen, G. Gutin, and A. Yeo, "When the greedy algorithm fails," *Discrete Optimization*, vol. 1, pp. 121-127, 2004.
 - [58] R. E. Schapire, "The Strength of Weak Learnability," *Mach. Learn.*, vol. 5, pp. 197-227, 1990.
-

-
- [59] L. G. Valiant, "A theory of the learnable," presented at the Proceedings of the sixteenth annual ACM symposium on Theory of computing, 1984.
- [60] J. Liu and S. Chen, "Resampling LDA/QR and PCA+LDA for Face Recognition
- AI 2005: Advances in Artificial Intelligence." vol. 3809, S. Zhang and R. Jarvis, Eds., ed: Springer Berlin / Heidelberg, 2005, pp. 1221-1224.
- [61] X. Lu and A. K. Jain, "Resampling for face recognition," presented at the Proceedings of the 4th international conference on Audio- and video-based biometric person authentication, Guildford, UK, 2003.
- [62] M. Most. (2003) Battle of the Biometrics. *Digital ID World Magazine*. 16-18.
- [63] A. Bendjebbour, *et al.*, "Multisensor image segmentation using Dempster-Shafer fusion in Markov fields context," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 39, pp. 1789-1798, 2001.
- [64] X. Chen, P. J. Flynn, and K. W. Bowyer, "IR and visible light face recognition," *Comput. Vis. Image Underst.*, vol. 99, pp. 332-358, 2005.
- [65] A. A. Ross, K. Nandakumar, and A. K. Jain, "Appendices," in *Handbook of Multibiometrics*, ed: Springer, 2006.
- [66] A. Tsymbal, M. Pechenizkiy, and P. Cunningham, "Diversity in search strategies for ensemble feature selection," *Information Fusion*, vol. 6, pp. 83-98, 2005.
- [67] A. Vezhnevets, "GML AdaBoost Matlab Toolbox," 0.3 ed. Moscow.
- [68] J. Grim, J. Kittler, P. Pudil, and P. Somol, "Information Analysis of Multiple Classifier Fusion," presented at the Proceedings of the Second International Workshop on Multiple Classifier Systems, 2001.
- [69] L. I. Kuncheva, "A Theoretical Study on Six Classifier Fusion Strategies," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, pp. 281-286, 2002.
- [70] L. I. Kuncheva, *Combining Pattern Classifiers: Methods and Algorithms*: Wiley, 2004.
- [71] J. Kittler, M. Hatef, R. P. W. Duin, and J. Matas, "On Combining Classifiers," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 20, pp. 226-239, 1998.
- [72] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Recognition*, 2nd ed.: John Wiley & Sons Inc, 2001.
-

-
- [73] B. Chapais and C. M. Berman, *Kinship and behavior in primates*. Oxford University Press, 2004.
- [74] P. G. Hepper, *Kin Recognition*. Cambridge University Press, 2005.
- [75] T. J. Bergman, J. C. Beehner, D. L. Cheney, and R. M. Seyfarth, "Hierarchical classification by rank and kinship in baboons," *Science*, vol. 302, pp. 1234-6, Nov 14 2003.
- [76] W. D. Hamilton, "The genetical evolution of social behaviour. I," *Journal of Theoretical Biology*, vol. 7, pp. 1-16, 1964.
- [77] J. R. Krebs and N. B. Davies, *Behavioral ecology*. Blackwell Science, 1997.
- [78] A. Alvergne, *et al.*, "Cross-cultural perceptions of facial resemblance between kin," *J Vis*, vol. 9, pp. 23 1-10, 2009.
- [79] L. DeBruine, B. Jones, A. Little, and D. Perrett, "Social Perception of Facial Resemblance in Humans," *Archives of Sexual Behavior*, vol. 37, pp. 64-77, 2008.
- [80] S. M. Platek, J. P. Keenan, and F. B. Mohamed, "Sex differences in the neural correlates of child facial resemblance: an event-related fMRI study," *Neuroimage*, vol. 25, pp. 1336-44, May 1 2005.
- [81] L. M. DeBruine, "Trustworthy but not lust-worthy: context-specific effects of facial resemblance," *Proc Biol Sci*, vol. 272, pp. 919-22, May 7 2005.
- [82] R. L. Burch and G. G. Gallup, "Perceptions of paternal resemblance predict family violence," *Evol Hum Behav*, vol. 21, pp. 429-435, Nov 2000.
- [83] S. M. Platek and S. M. Kemp, "Is family special to the brain? An event-related fMRI study of familiar, familial, and self-face recognition," *Neuropsychologia*, vol. 47, pp. 849-58, Feb 2009.
- [84] P. J. Phillips, M. Hyeonjoon, S. A. Rizvi, and P. J. Rauss, "The FERET evaluation methodology for face-recognition algorithms," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 22, pp. 1090-1104, 2000.
- [85] X. Tan, S. Chen, Z.-H. Zhou, and F. Zhang, "Face recognition from a single image per person: A survey," *Pattern Recogn.*, vol. 39, pp. 1725-1745, 2006.
- [86] T. Cover and P. Hart, "Nearest neighbor pattern classification," *Information Theory, IEEE Transactions on*, vol. 13, pp. 21-27, 1967.
- [87] L. Fayin, H. Wechsler, and M. Tistarelli, "Robust fusion using boosting and transduction for component-based face recognition," in *Control, Automation,*
-

- Robotics and Vision, 2008. ICARCV 2008. 10th International Conference on*, 2008, pp. 434-439.
- [88] C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20, pp. 273-297, 1995.
- [89] K. Nandakumar, C. Yi, S. C. Dass, and A. K. Jain, "Likelihood Ratio-Based Biometric Score Fusion," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 30, pp. 342-347, 2008.
- [90] Available: <http://dictionary.reference.com/browse/similarity>
- [91] R. McDonnell and A. McNamara, "Application of the Golden Ratio to 3D Facial Models," in *the 2003 Eurographics Ireland Workshop*, Ireland, 2003, pp. 87-103.
- [92] K. Goebel and W. Yan, "Choosing Classifiers for Decision Fusion," in *Proc. of the Seventh International Conference on Information Fusion*, 2004, pp. 563-568.
- [93] B. Gökberk, M. O. Irfanoglu, L. Akarun, and E. Alpaydn, "Selection of Location, Frequency, and Orientation Parameters of 2D Gabor Wavelets for Face Recognition," in *Advanced Studies in Biometrics*. vol. 3161/2005, ed Berlin: Springer, 2005, pp. 138-146.
- [94] A. K. Jain, B. Klare, and U. Park, "Face recognition: Some challenges in forensics," in *Automatic Face & Gesture Recognition and Workshops (FG 2011), 2011 IEEE International Conference on*, 2011, pp. 726-733.
- [95] T. Ojala, M. Pietikäinen, and D. Harwood, "A comparative study of texture measures with classification based on featured distributions," *Pattern Recognition*, vol. 29, pp. 51-59, 1996.
- [96] S. Krishna, J. Black, and S. Panchanathan, "Using Genetic Algorithms to Find Person-Specific Gabor Feature Detectors for Face Indexing and Recognition," in *Advances in Biometrics*. vol. 3832, D. Zhang and A. Jain, Eds., ed: Springer Berlin / Heidelberg, 2005, pp. 182-191.
- [97] Z. Sun, G. Bebis, and R. Miller, "Evolutionary Gabor Filter Optimization with Application to Vehicle Detection," presented at the Proceedings of the Third IEEE International Conference on Data Mining, 2003.
- [98] T. Jabid, M. H. Kabir, and C. Oksam, "Local Directional Pattern (LDP) for face recognition," in *Consumer Electronics (ICCE), 2010 Digest of Technical Papers International Conference on*, 2010, pp. 329-330.

-
- [99] D. He and N. Cercone, "Local Triplet Pattern for Content-Based Image Retrieval," presented at the Proceedings of the 6th International Conference on Image Analysis and Recognition, Halifax, Nova Scotia, Canada, 2009.
- [100] B. Gökberk, M. O. İrfanoğlu, L. Akaruna, and E. Alpaydın, "Learning the best subset of local features for face recognition," *Pattern Recognition*, vol. 40, May 2007.
- [101] Y. LeCun, L. Bottou, P. Haffner, and P. Howard, "Djvu: a compression method for distributing scanned documents in color over the internet," in *In Color 6. IST*, ed, 1998.
- [102] J. W. Mann and G. D. Smith, "A comparison of heuristics for telecommunications traffic routing," in *Modern Heuristic Search Methods*, V. J. Rayward-smith, *et al.*, Eds., ed Chichester: Wiley, 1996, pp. 235-254.
- [103] H. Mühlenbein and D. Schlierkamp-Voosen, "Predictive models for the breeder genetic algorithm i. continuous parameter optimization," *Evolutionary Computation*, vol. 1, pp. 25-49, 1993.
- [104] T. Ahonen, A. Hadid, and M. Pietikäinen, "Face Recognition with Local Binary Patterns." vol. 3021, T. Pajdla and J. Matas, Eds., ed: Springer Berlin / Heidelberg, 2004, pp. 469-481.
- [105] T. Ojala, *et al.*, "Outex - New Framework for Empirical Evaluation of Texture Analysis Algorithms," presented at the Proceedings of the 16 th International Conference on Pattern Recognition (ICPR'02) Volume 1 - Volume 1, 2002.
- [106] B. Caputo, E. Hayman, and P. Mallikarjuna, "Class-Specific Material Categorisation," presented at the Proceedings of the Tenth IEEE International Conference on Computer Vision - Volume 2, 2005.
- [107] A. Alvergne, C. Faurie, and M. Raymond, "Differential facial resemblance of young children to their parents: who do children look like more?," *Evolution and Human Behavior*, vol. 28, pp. 135-144, 2007.
- [108] M. Ghahramani, W. Y. Yau, and E. K. Teoh, "Family facial patch resemblance extraction," presented at the Proceedings of the 10th Asian conference on Computer vision - Volume Part II, Queenstown, New Zealand, 2011.
- [109] L. M. DeBruine, *et al.*, "Kin recognition signals in adult faces," *Vision Res*, vol. 49, pp. 38-43, Jan 2009.
-

-
- [110] N. J. Christenfeld and E. A. Hill, "Whose baby are you?," *Nature*, vol. 378, p. 669, Dec 14 1995.
 - [111] S. Bredart and R. M. French, *Do Babies Resemble Their Fathers More Than Their Mothers? A Failure To Replicate Christenfeld And Hill (1995)* vol. 20. New-York, NY, ETATS-UNIS: Elsevier, 1999.
 - [112] M. Pagel, "Desperately concealing father: a theory of parent–infant resemblance," *Animal Behaviour*, vol. 53, pp. 973-981, 1997.
 - [113] P. Bressan, "Why babies look like their daddies: paternity uncertainty and the evolution of self-deception in evaluating family resemblance," *Acta ethologica*, vol. 4, pp. 113-118, 2002.
 - [114] C. A. Meissner and J. C. Brigham, "Thirty Years of Investigating the Own-Race Bias in Memory for Faces: A Meta-Analytic Review," *Psychology, Public Policy, and Law*, vol. 7, pp. 3-35, 2001.
 - [115] K. A. Toh, J. Xudong, and Y. Wei-Yun, "Exploiting global and local decisions for multimodal biometrics verification," *Signal Processing, IEEE Transactions on*, vol. 52, pp. 3059-3072, 2004.
 - [116] S. Prabhakar and A. K. Jain, *Decision-level fusion in fingerprint verification* vol. 35. Kidlington, ROYAUME-UNI: Elsevier, 2002.
 - [117] B. Ulery, *et al.*, "Studies of Biometric Fusion," NIST, Technical Report IR 7346, 2006.
 - [118] M. A. T. Figueiredo and A. K. Jain, "Unsupervised Learning of Finite Mixture Models," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, pp. 381-396, 2002.
 - [119] G. Griffin, A. Holub, and P. Perona, "The Caltech-256," Caltech Technical Report.
 - [120] T. Peng. (2010). *Detect circles with various radii in grayscale image via Hough Transform.* Available: <http://www.mathworks.com/matlabcentral/fileexchange/9168>