TOWARDS FACIAL ASYMMETRY BASED FACE RECOGNITION

By

Muhammad Said
PE101005

A thesis submitted to the
Department of Electrical Engineering
In partial fulfillment of the requirements for the degree of
PhD IN ELECTRONIC ENGINEERING

Faculty of Engineering
Capital University of Science & Technology
Islamabad
September, 2016

Dedicated to my beloved ones

ACKNOWLEDGMENTS

In the name of Allah, most gracious most merciful, the creator of all that exists. I would like to express my sincere gratitude to my supervisor Prof. Dr. Imtiaz A. Taj who has always been a fatherly figure for me for the past five years. I am indebted for his valuable guidance on both course work and research matters. I have had a very pleasant stay at Vision and Pattern Recognition Research Laboratory. I enjoyed my collaborations and discussions with my research group members. I am especially beholden to my close collaborators, the research group members for all the fruitful scientific discussions. This work would not have been possible without this group. I would also like to thank my advisor and administration for awarding scholarship during my PhD here at Capital University of Science and Technology, Islamabad.

ABSTRACT

Face recognition, as an active research area over the past three decades, still poses many challenges. Recognition of age-separated face images (age invariant face recognition) based on facial asymmetry is one of such challenges. Successful solutions to this recognition paradigm would allow the facial photographs to be matched against face images with temporal variations. Facial asymmetry, which refers to non-correspondence in shape, size, and arrangement of facial landmarks on both sides of the face, is an intrinsic recognition-specific facial feature used for face recognition task. The contributions of this dissertation are focused on recognition of age-separated face images using facial asymmetry. We introduce to use a feature description scheme suitable to represent facial asymmetry. The introduced feature description is adaptable to recognize age-separated face images and extract demographic information such as age group, gender, and race from a given face image. Based on the introduced feature description, this dissertation offers the following three main contributions to recognize age-separated face images.

The first contribution is a matching-scores space based approach to recognize age separated face images. In the proposed framework, matching scores of holistic, local, and asymmetric facial features are combined in a matching-score space (MSS) with Support Vector Machine (SVM) as a classifier to separate genuine and imposter classes. Experimental results on three publically available benchmark facial aging databases show the efficacy of proposed approach compared to some existing state-of-the-art approaches.

The second contribution is focused on the role of facial asymmetry based age group estimation in recognizing age-separated face images. We provide a hierarchical approach to perform age group estimation task. The role of various asymmetric facial regions in recognizing age-separated face images of different age groups is investigated. We integrate the knowledge learned from age group estimation into face recognition algorithm to enhance the recognition performance of age-separated face images. The viability of this approach is demonstrated on two benchmark facial aging databases. The experimental results suggest that integration of age group estimates into face recognition algorithm enhances the recognition performance of age separated face images, considerably.

The third contribution is examination of the role of facial asymmetry in demographic estimation (i.e. age group, gender, and race) of a query face image in a face recognition system. The role of different asymmetric facial regions in recognizing face images with different demographic attributes is presented. We integrate the demographic estimates into a face recognition algorithm to enhance the recognition accuracy of age-separated face images. Experiments are conducted on benchmark facial aging databases to validate the performance of proposed approach. The experimental results suggest that proposed approach is more adaptable to recognize age-separated face images compared to some existing state-of-the-art methods.

LIST OF PUBLICATIONS

Journal Publications

1. Muhammad Sajid, Imtiaz A Taj, Usama I Bajwa, Naeem I Ratyal, "The role of Facial Asymmetry in Recognizing Age-Separated Face Images," *Computers and Electrical Engineering* (2016), doi: http://dx.doi.org/10.1016/j.compeleceng.2016.01.001

2. Naeem Iqbal Ratyal, Imtiaz A Taj, Usama I Bajwa, Muhammad Sajid, "3D fcae recognition based on pose and expression invariant alignment ," *Computers and Electrical Engineering,* vol. 46, pp. 241-255, 2015.

3. Muhammad Sajid, Imtiaz A Taj, Usama I Bajwa, Naeem I Ratyal, "Facial Asymmetry based Age Group Estimation: Role towards Recognizing Age-Separated Face Images ," submitted to *Neural Computing and Applications (under Review).*

4. Muhammad Sajid, Imtiaz A Taj, Usama I Bajwa, Naeem I Ratyal, "The Role of Facial Asymmetry based Demographic Estimation in Recognizing Age-Separated Face Images ," submitted to *IEEE Transactions on Information Forensics and Security (under Review).*

5. Naeem I Ratyal, Imtiaz A Taj, Usama I Bajwa, Muhammad Sajid, "Automatic Multi-View Pose and Expression Invariant 3D Face Recognition ," submitted to *Computers and Electrical Engineering (under Review).*

Conference Publications

1. Naeem I Ratyal, Imtiaz A Taj, Usama I Bajwa, Muhammad Sajid, Mirza Jabbar A Baig and Faisal M Butt, "3D Face Recognition based on Region Ensemble and Hybrid Features", *in Proceedings of the International Conference on Computing, Electronic and Electrical Engineering (ICE Cube),* April 11-12, 2016.

TABLE OF CONTENTS

Chapter 3

Chapter 4

Chapter 5

Chapter 6

Chapter 8

LIST OF FIGURES

LIST OF TABLES

LIST OF ACRONYMS

AAM	Active Appearance Model
AFs	Attribute Filters
APs	Attribute Profiles
ASM	Active Shape Model
BIF	Biologically Inspired Features
CCA	Canonical correlation analysis
CMC	Cumulative Match Characteristics
DOG	Difference of Gaussian
DSAF	Densely Sampled Asymmetric Features
FAR	False Acceptance Rate
FERET	Facial Recognition Technology
FRVT	Face Recognition Vendor Test
GOP	Gradient Orientation Pyramid
ICA	Independent Component Analysis
LBP	Local Binary Pattern
LDA	Linear Discriminant Analysis
NIST	National Institute of Standards and Technology
PCA	Principal Component Analysis
ROC	Receiver Operating Characteristics
SIFT	Scale Invariant Feature Transform
SVM	Support Vector Machine

Chapter 1

INTRODUCTION

1.1 Overview

Face recognition is a topic of recent research that uses computer algorithms to determine the similarity between two face images [1]. Recognition of face images finds numerous applications in public security, criminal investigations, photo-ID issuance, photo indexing, and web access.

In past three decades a lot of advancements have been made in the field of face recognition with an ample amount of related literature. It has been observed psychologically that a number of quite complex factors are involved in recognizing known and unknown faces. In case of known faces, recognition seems to be independent of variations like pose and expressions, whereas unknown faces are affected by such variations. In biology, it has been observed that brain takes the faces as very different from other objects. For example, standard studies in this direction showed that if the face image is vertically inverted then recognition is reduced drastically. In a pioneering work [2] it was verified that there is a special brain area called as fusiform that is responsible for recognizing human faces. However some recent revisions [3] suggest that the process of faces recognition and other objects relies on same mechanisms. Scientists are using similar ideas to be implemented in machines to differentiate faces from other objects.

The progress made in face recognition technology over the past two and half decades, has been benchmarked by National Institute of Standards and Technology (NIST) [4]. The most recent Face Recognition Vendor Test (FRVT) reports the accuracy of leading commercial “off-the-shelf” algorithms between 80% and 97%, depending on the face recognition algorithms. One can assume that error rates have been dropped significantly, thus face recognition has become a largely a solved problem. However face recognition of real world images still poses many challenges. The reduction in error rates shown in

Figure 1.1 is for face images acquired under controlled conditions. However, the face recognition performance deteriorates when certain changes like illumination, expressions, head-pose, and aging variations are introduced [5]. Some example face images exhibiting these variations are shown in Figure 1.2. Other aspects such as image quality, segmentation, and compression artifacts also contribute to the recognition accuracy of face images [6].

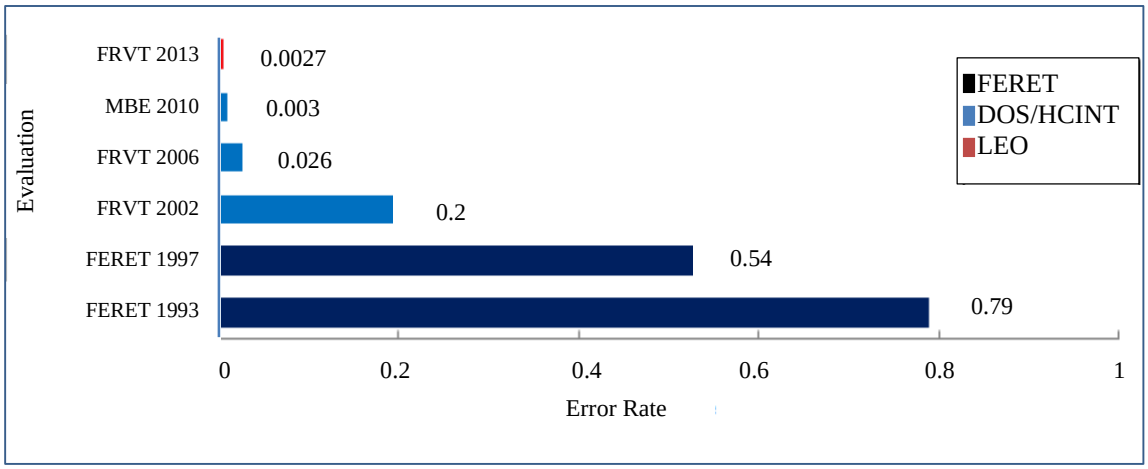


Figure 1.1: Reduction in error rates for state-of-the-art algorithms benchmarked by NIST

(a) (b) (c) (d)

Age 50 54 54 56 55 57

Figure 1.2: Facial variations: Top row: (a) neutral, (b) pose, (c) illumination, and (d) expression. Bottom row: Example age-separated face image pairs showing intra-subject temporal variations (age in years)

One of the most challenging tasks of the modern face recognition algorithms is matching age-separated face images (i.e., age invariant face recognition). Developing algorithms to recognize age-separated face images can prove to be beneficial in certain applications such as border control, locating missing persons, and child trafficking.

1.2 Face as a Biometric used for Verification or Identification Task

Face, like other biometrics (for example, finger print and iris) can be used for verification or identification tasks, as described below.

(i) Face verification: The verification task aims at deciding if the probe image and the requested identity are matched or not.

(ii) Identification: Identification task aims at first deciding the identity of the query face image and then comparing it with gallery images with pre-known identities.

Face verification and identification tasks performed by a general face recognition system are illustrated in Figure 1.3.

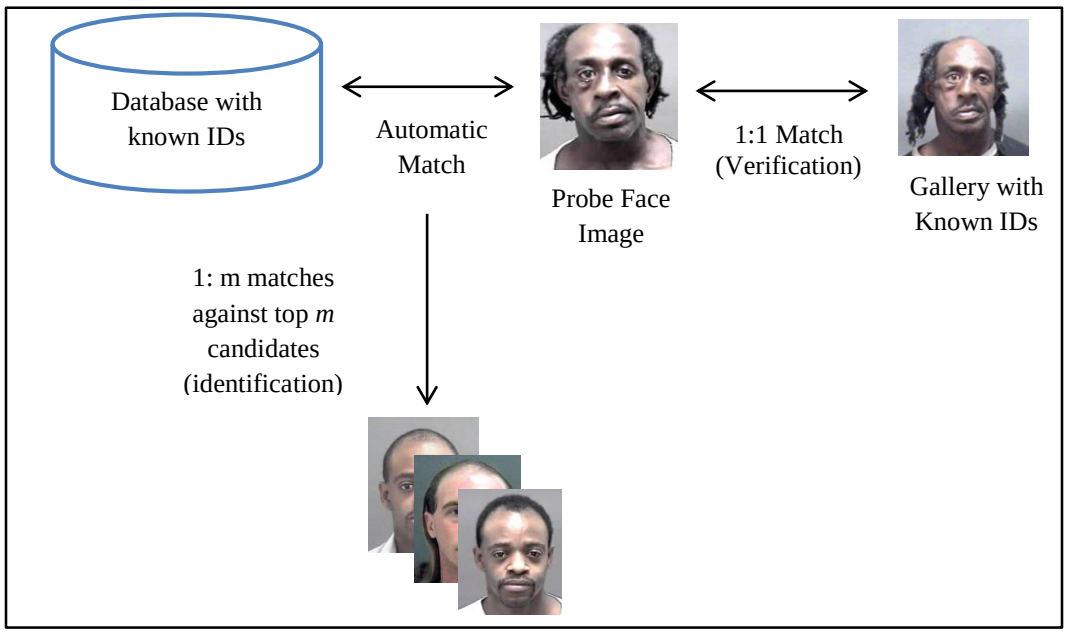


Figure 1.3: Main tasks of a general face recognition system

1.3 A General Face Recognition System

A general face recognition system consists of multiple components as shown in Figure 1.4. The main components of a general face recognition system include face detection and face recognition. The later component can be further divided into following three parts.

(i) Face normalization

(ii) Feature extraction

(iii) Classification

A face can be detected by using a face detector. A comprehensive survey on face detection has been given in [7].

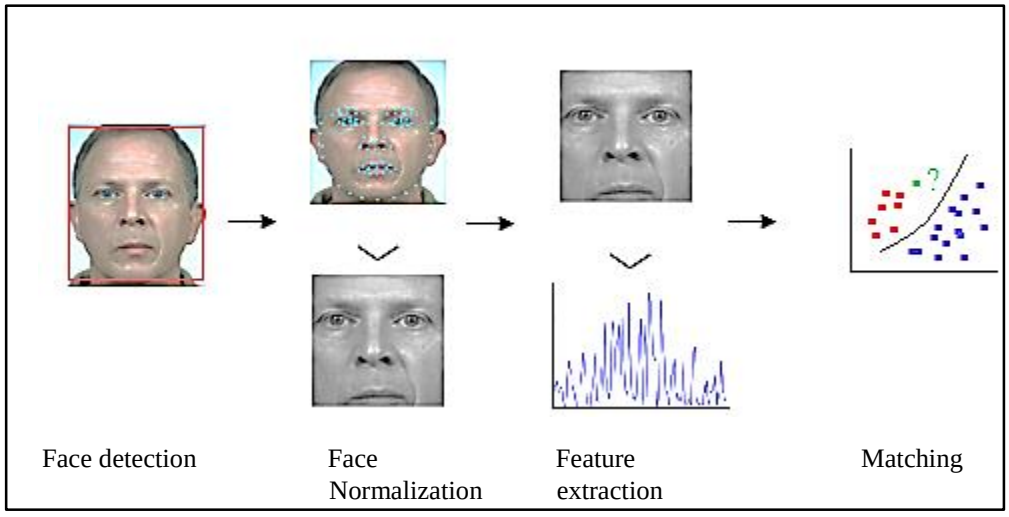


Figure 1.4: A general face recognition system

Face normalization is carried out prior to recognition task and it generally involves two variants: (i) the illumination normalization, and (ii) geometric normalization. Illumination normalization is employed to limit any undesirable lighting variations and shadows. Geometric normalization, also called alignment, involves detection of some specific points on the face, including facial landmarks such as eyes centers, tip of the nose, and facial contour etc. These landmark points are then secured to predefined positions for all

images. For example, the eyes-centers, and facial mid line are made to be fixed at the similar pixel coordinates in all the images to be aligned. Feature extraction stage provides the evocative information to be used later in the classification stage where a query face image is matched with one of the gallery images using facial features.

1.4 Problem Statement

Figure 1.5 shows the impact of facial aging on recognition performance for a commercial off the shelf (COTS) face recognition system [8]. The performance has been measured as a function of time lapse between probe and gallery images, on a mug shot database consisting of 94631 face images of 28031 individual subjects. The performance degradation illustrates the complex nature of face recognition of age-separated face images. By comparison, it is observed that when the age gap between gallery and probe images is large, the recognition performance decreases substantially. This performance degradation suggests that there is vast scope of improvement in recognizing age-separated face images algorithms.

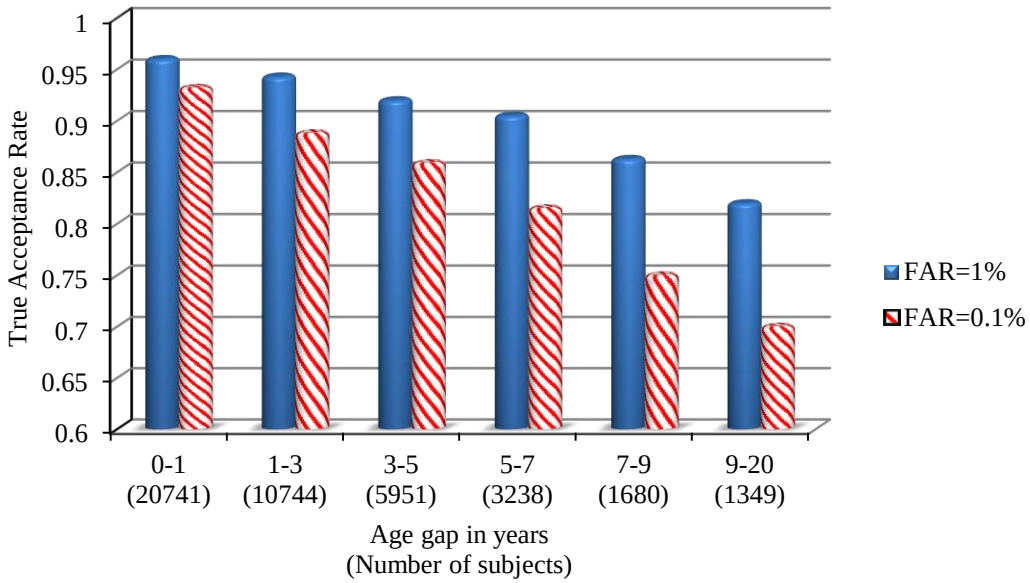


Figure 1.5: The decrease in face recognition accuracy with increasing age for a state-of-the-art algorithm [8]

Published methods to recognize age-separated face images are limited. These methods use either holistic features such as principal component analysis (PCA) (eigen faces [9]) and linear discriminat analysis (LDA) (Fisherfaces [10]), or local facial features, such as local binary patterns (LBP) [11] and scale invariant feature transform (SIFT) [12] to recognize age separated face images. While existing approaches have achieved remarkable performance in recognizing age-separated face images, yet there are two main limitations:

(i) There is lack of facial features evaluation along with temporal variations.

(ii) Most of the existing methods such as [11], [13], [14] rely on a single set of features and hence lack the discriminative information. For example appearance based facial features exhibit poor performance in discriminating age-separated face images due to their dependence on raw pixel-intensity [15].

Local features, though used in many age invariant face recognition methods, can be manipulated by certain age groups due to artificial extrinsic facial variations (for example, by facial make up [16]). In order to improve recognition performance across aging, it is imperative to introduce such intrinsic facial features that can be used as a strong indicator of age.

Facial asymmetry, which refers to non-correspondence in shape, size, and arrangement of facial landmarks on both sides of the face, is common in humans, even in young healthy subjects [17]. Being an intrinsic facial characteristic and strong indicator of age, facial asymmetry can be used to learn important characteristics such as features to classify face images into different age groups and to recognize age-separated face images. Figure 1.6 illustrates the common steps utilized by recognition of age-separated face images algorithms using asymmetric facial features. A face recognition model is trained using asymmetric facial features and matching is performed based on distance calculation of asymmetric facial feature of a query face image with that of training face images.

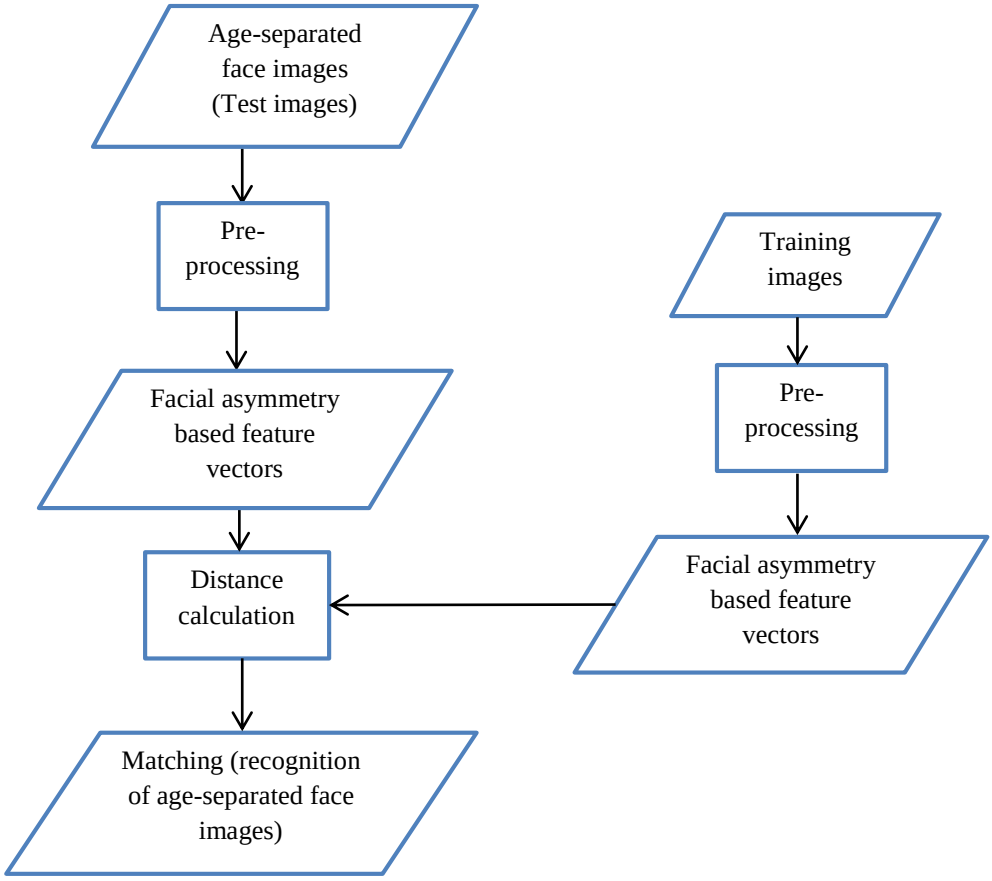


Figure 1.6: The common steps utilized by recognition of age-separated face images algorithms

1.5 Purpose of Research

Some main tasks of the research presented in this dissertation are listed below.

(i) To review the different stages of face recognition systems including pre-processing, feature extraction, and matching protocols.

(ii) To review different factors contributing facial aging process including facial asymmetry.

(iii) To recognize age-separated face images using facial asymmetry. For this purpose, three separate studies are presented as described below.

(1) The first study [18] involves

- Measurement and evaluation of facial asymmetry with temporal variations.
- Recognition of age-separated face images using asymmetric facial features along with existing facial features in proposed matching-scores space (MSS).

(2) The second study [19] focuses

- Facial asymmetry based age group estimation.
- Role of different asymmetric facial regions in recognizing age-separated face images.
- Integration of knowledge learned from age group estimation into a face recognition algorithm to enhance recognition performance.

(3) The third study [20] seeks

- Role of facial asymmetry in demographic estimation (age group, gender, and race).
- Integration of demographic estimates into a face recognition algorithm to improve recognition performance of age-separated face images.

1.6 Thesis Organization

The thesis consists of eight chapters organized as follows.

Chapter 2 presents a detailed literature survey on face recognition of age separated face images, age group estimation, and demographic estimation. Literature survey on facial asymmetry is also presented.

In Chapter 3, we present a comprehensive account of facial feature representation and feature extraction.

In Chapter 4, introduced feature representation scheme is presented.

In Chapter 5, we present a score pace based solution to the problem of recognizing age separated face images, using existing and proposed asymmetric facial features. An account of facial asymmetry measurement and evaluation is also presented.

In Chapter 6, we present facial asymmetry based age group estimation. Role of different asymmetric facial regions to recognize age-separated face images is presented. Integration of knowledge learned from age group estimation with algorithm to recognize age-separated face images is also presented.

Chapter 7 presents a comprehensive account of facial asymmetry based demographic estimation and its role in recognizing age-separated face images.

Chapter 8 concludes the findings of this dissertation and suggests future research directions.

Chapter 2

LITERATURE REVIEW

The goal of this chapter is to provide an insight into background and pre-requisites of face recognition. First an overview of specifics of a typical face recognition system with face detection and face normalization is highlighted, and then a brief overview of most popular approaches in face recognition systems is presented. Finally a description of face images databases is presented.

2.1 Face Recognition: Prerequisites

Most of the face recognition algorithms perform matching between a given pair of face images based on recognition task consisting of the following steps.

(i) Face detection
(ii) Face normalization
(iii) Feature representation
(iv) Feature extraction
(v) Matching

An overview of each of the above mentioned steps is given in the following subsections.

2.1.1 Face Detection

Often viewed as a pre-processing step, face detection involves detecting the presence of a face in a digital image. Face detector proposed in [21] is used as a precedent for all modern face detectors and serves as a baseline of state-of-the-art performance [22].

2.1.2 Face Normalization

Face normalization is an essential step of any face recognition system and it generally involves geometric and illumination normalizations.

Geometric normalization is typically performed by detecting facial landmarks, generally by using eyes coordinates. Most commonly used geometric normalization methods are described below.

(i) Normalization using 3D morphable model involves detection of a more verbose set of facial landmarks, such as eyes centers, nose tip, and facial outline to fix the distance between corresponding landmarks of all face images to be aligned.

(ii) Normalization using 2D affine transformation involves detection of eyes coordinates. Planer rotation and scaling is then performed to fix the distance and angle between two eyes for all face images to be aligned.

Facial landmarks can generally be detected by using Active Appearance Model [23] (AAM), or Active Shape Model (ASM) [24]. Face++ [25] is another recent state-of-the-art method based on deep face representation for accurate detection of facial landmarks. Figure 2.1 shows the examples of geometric face normalization.

Illumination normalization seeks corrections for unwanted lighting variations. A number of methods have been proposed to perform these corrections such as histogram equalization and Difference of Gaussian (DoG) filters [26].

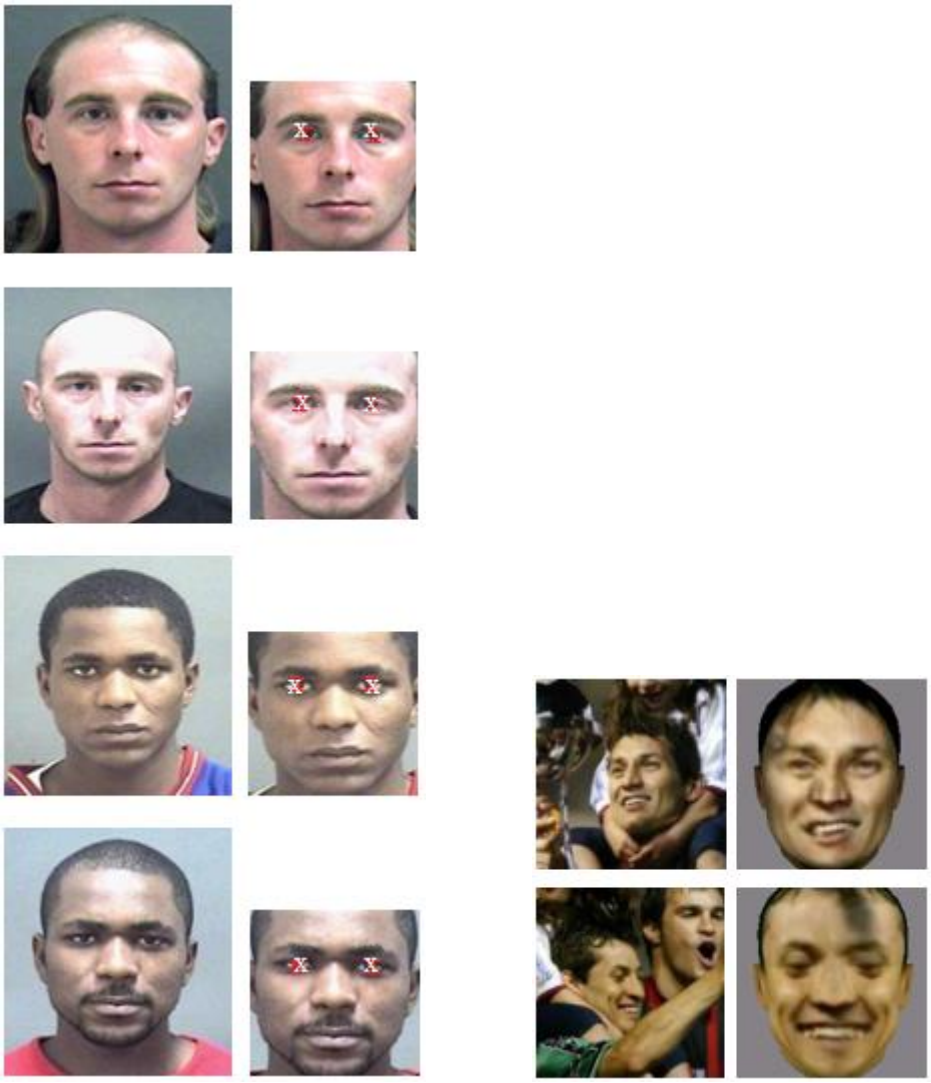
2.1.3 Feature Representation

Facial feature representation stage represents different facial characteristics in the form of a feature descriptor vectors. Such feature vectors may contain image information in the form of pixels or other measurements such as distances between facial landmarks or even more complex features such as convolution of face images with Gabor filters. In order to develop an organization, taxonomy of facial features is given in [15].

2.1.4 Feature Extraction

Feature extraction on the other hand focuses on feature combinations which project the original feature descriptors into a feature space with improved class separation. Generally, feature extraction can be carried out either as a holistic representation or a

local representation. A detailed account of facial feature representation and extraction methods reported in literature is presented in Chapter 3.



(a) (b)

Figure 2.1: Geometric face normalization (a) Face images before normalization (left column) and face images after 2D affine transformation based on eyes coordinates (right column) (images are taken from MORPH database [27]). (b) Face images before normalization (left column) and face images normalized using 3D morphable model (right column) (images are taken from [28])

2.1.5 Matching

In matching stage, feature vectors computed through feature extraction stage are used to compute a measure of similarity (or dissimilarity) between two given face images.

In a typical face recognition system, face images are enrolled into the system as gallery data and face image of a given test subject (probe image) is matched to the gallery data using one of the following two schemes.

(i) A one-to-one matching, also called face verification

(ii) A one-to-many matching, also called face identification.

Nearest neighbor (NN) classification algorithm [29] is the most simple among matching techniques. A probe image is matched with gallery face image with minimum distance, such as MahCos or Euclidean distance metric. K-nearest neighbor (k-NN) is a density estimation based method in which the probability of a test point t' residing within a volume V centered at some point t is approximated by proportions of samples S residing within V, as shown in equation (2.1).

$$p(t) = \frac{K}{nV} \tag{2.1}$$

The k-NN classifier assigns a point t to a particular class based on a majority vote among the classes of the k nearest training points to t.

In face recognition algorithms, often the feature vectors are used to calculate matching scores to find similarity or dissimilarity between a pair of face images. This notion can easily leverage a host of binary classification algorithms from machine learning and pattern recognition literature by creating genuine and imposter matching scores. Genuine scores represent a true match while imposter scores represent a false match between a given pair of face images. Binary classifiers such as Support Vector Machines (SVMs) [30] can be used to separate genuine and imposter scores.

SVMs facilitate the classification between two classes with a hyper plane by maximizing the distance to the closest points in the training. In binary classification problem, a set of training matching scores belonging to two different classes, $(u_1, v_1), (u_2, v_2), \ldots\ldots, (u_i, v_i)$

where $u_i \in R_n$ and$v_i \in \{1, -1\}$, can be separated by a hyper plane$w_u + b = 0$. Hence if u_i is a genuine score, then $v_i = 1$, otherwise $v_i = -1$. The maximum margin defined in this case is termed as Optimal Separating Hyper-plane (OSH). The OSH aims at minimizing the objective function as shown in equation (2.2).

$$\psi(w, \xi) = \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{j}\xi_i \tag{2.2}$$

subject to following constraints

$$V_i = [(w, u_i] \geq 1 - \xi_i, \xi_i \geq 0, i = 1, \dots, j$$

where C is regularization parameter, which is used to achieve trade-off between complexities of the Support Vector Machine and ξ_i represent slack variables used to penalize errors when data is not linearly separable.The sign of OSH decision surface function φ can be used to classify a test point u as shown in equation (2.3).

$$\varphi(u) = \sum_{i=1}^{j} v_i \beta_i K(u, u_i) + b \tag{2.3}$$

where $\beta_i \geq 0$ are defined as Langrangian multipliers and b is solution of optimization problem as defined in the above mentioned formulation. $K(u, u_i)$ is kernel trick used to transform data to a higher dimensional space where it is more separable. One of the most popular choices is the radial basis function RBF kernel, defined as shown in equation (2.4).

$$K(u, v) = exp(\frac{-|u - v|^2}{\delta^2}) \tag{2.4}$$

where δ is a positive constant, which defines the width of the kernel.

Certain fusion strategies, such those employed in [31] may effectively be used to enhance matching results, for example, multiple facial features (such as holistic and local features), or multiple views of a face (such as frontal and profile facial views) can be consolidated to achieve better classification results.

2.2 Face Recognition: Literature Review

This chapter gives an overview of face recognition algorithms. The next chapter gives a detailed account about the future representation and feature extraction algorithms employed in face recognition algorithms. Detailed survey reports on face recognition methods are presented in [32], [33] which are good resource for gaining knowledge about existing state-of-the-art methods in face recognition paradigm. The face recognition algorithms include feature analysis based methods, such as Active Appearance Models (AAMs), Active Shape Models (ASMs) and classifier intelligence based methods such as Elastic Bunch Graph Matching (EBGM), and Support Vector Machines (SVMs).

Face recognition algorithms presented in this chapter are broadly grouped into three categories, general face recognition algorithms including appearance and model based methods, Algorithms for facial aging analysis including recognition of age-separated face images, and demographic estimation and facial asymmetry based face recognition algorithms, which is the topic under current research.

A summarized description of these face recognition algorithms is given in the following subsections.

2.2.1 General Face Recognition Algorithms

General face recognition algorithms can be classified as holistic or appearance based methods and model based methods as described below.

Holistic or appearance based methods: In these methods, raw intensity images are used to represent a given face image. A face is assumed as an image point residing in some high dimensional vector space. A number of appearance based strategies make use of statistical based methods to analyze the distribution of the image vectors in the vector space and develop an effective representation accordingly. Any match between a probe and gallery images is then determined in the corresponding feature space. Principal Component Analysis (PCA), is a general way of calculating the subspace having the highest variance of data. In this regard, eigen face method is a famous appearance based

method [9] which employs PCA to construct a subspace. The highest amount of variance is determined in terms of orthogonal directions which are determined through PCA. A subspace covered by these directions is then employed to project the data. Practically, the principal component axes are eigen vectors of the covariance matrix of the data. The corresponding eigen values indicate the proportion of variance of the data projections along each direction. Eigen analysis is then used to find the largest eigen vectors. By projecting the face image into the coordinate system defined by the eigen faces gives its representation in PCA subspace. PCA tends to maximize the overall data variance. However, PCA suffer from the fact that it tends to increase within class scatter than that of between class scatter due to variations like pose and illumination.

In order to overcome this limitation of PCA, Linear Discriminant Analysis (LDA) can be used to enhance recognition performance. In case of LDA, class-specific information is used to construct a subspace in which the class representation is optimized. LDA employs Fisher face method [10]. Fisher face algorithm is based on Fisher Linear Discrimanat (FLD) which employs class-specific information. In training, the images are divided into corresponding classes, based on different statistics. Then the techniques similar to those used with eigen face algorithm is applied. LDA tends to reduce within-class scatter compared to between-class scatter, resulting into enhanced recognition performance compared to PCA. Although LDA can outperform PCA, yet it requires a larger training set to achieve a good generalization, which is itself a challenge in face recognition paradigm.

Application of holistic methods (PCA and LDA) to a sample face image is shown in Figure 2.2, where the top row shows an example input face image, 2nd row shows the eigen faces corresponding to eight largest eigen values, while 3rd row shows the corresponding Fisher faces and the bottom row shows the reconstructed faces from PCA and LDA, respectively.

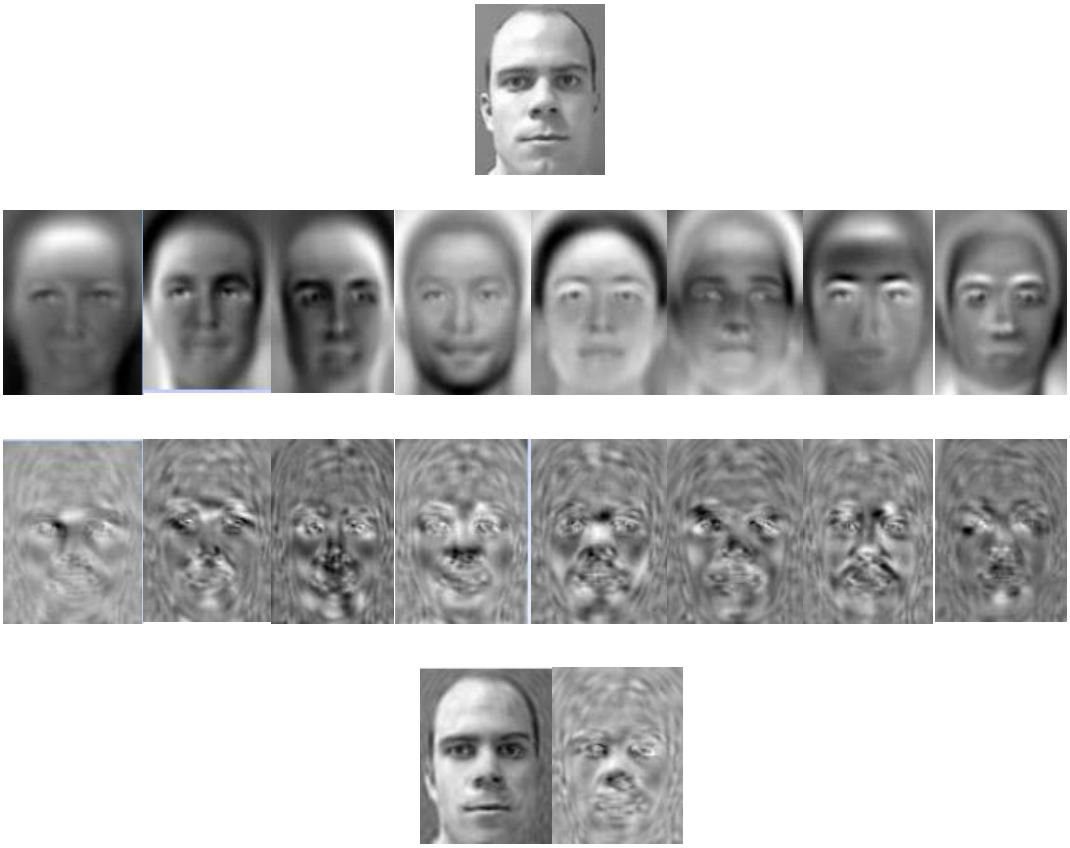


Figure 2.2: Example of holistic features: (top row) an example input face image, (2nd row) top eight eigen faces, (3rd row) corresponding Fisher faces, (bottom row) reconstructed faces from PCA and LDA

In addition to linear techniques such as PCA and LDA, non-linear techniques also exist. Nonlinear strategies involve, kernel based methods which can effectively be employed to in solving recognition problems. In nonlinear techniques, the sample data is transformed into some higher dimensional feature space. LDA or PCA analysis can subsequently be performed as described earlier. KPCA and GDA [34] are examples of nonlinear techniques which can outperform the PCA and LDA methods.

Model based methods: These approaches tend to build a human face model, which in turn capable of taking into account the facial variations like pose and expressions. To design such a model, a prior familiarity of a given face image is used. Feature based matching is an example of this approach e.g., Elastic Bunch Graph Matching (EBGM)

[35]. Similarly, Active Appearance and Shape Models consider both, the shape and texture to develop a 2D morphable facial model [36], which is then used to learn the face variations accordingly. These model based approaches are described below.

(i) Elastic Bunch Graph Matching (EBGM): In this strategy a face is represented by a labeled graph. The graph is a rectangular grid placed on the image. In the grid, nodes are characterized with jets representing the retorts of Gabor filters in various directions and spatial frequencies. The edges of the grid represent the distances. Matching is performed by comparing the graph of a probe image to that of a gallery image. The goal of Elastic Graph Matching is to extract a graph based on fiducial points (see Figure 2.3) from query face image and a graph similarity is maximized between query and gallery face image.

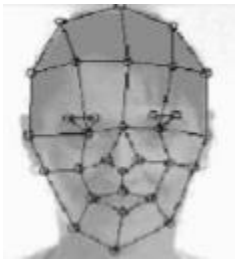


Figure 2.3: Graphs extracted from fiducial points for a sample face images [35]

The similarity $S_{\emptyset}$ between two jets is defined as shown in equation (2.5).

$$S_{\emptyset}(i,i') = \frac{\sum_j m_i p_i' \cos(\emptyset_i - \emptyset_i' - \overrightarrow{dk_i})}{\sqrt{\sum_i m_i^2 \sum_i p_i'^2}} \tag{2.5}$$

where m_i and p_i' are magnitude and phase of Gabor coefficients in the i^{th} jet, respectively, d is displacement between jets and k represents orientation and wavelength of Gabor wavelet kernels.

In graph matching approach, the graph is capable of rotation, translation, and scaling, thus compensating different variations in face images.

(ii) Active Appearance and Shape Model: A statistical based model which combines active shape model (ASM) with a model of the appearance variations [37]. This combination takes place in a shape normalized frame. In case of active appearance and shape model matching is performed by calculating the difference between the probe image and a synthesized model.

In order to construct appearance model, facial landmark positions are used to construct a vector $v = (x_1, y_1, \dots, x_n y_n)^T$ describing the shape of a face image, where (x_i, y_i) represent the 2D coordinates of i^{th} facial landmark. Principal Component Analysis (PCA) is performed on normalized shape vectors to construct the face shape model represented as shown in equation (2.6).

$$v = \bar{v} + O_v p_v \tag{2.6}$$

where v is a shape vector, $\bar{v}$ is mean shape, O_v is set of orthogonal mode of shape variations and p_v is set of shape parameters.In order to construct appearance model, intensity of sampled pixels in wrapped imag is used to construct a vector and then PCA is used to construct a linear model as shown in equation (2.7).

$$a = \bar{a} + O_a p_a \tag{2.7}$$

where a is appearance vector, $\bar{a}$ is mean appearance vector, O_a is set of orthogonal mode of gray-level variations and p_a is gray-level model parameters.

The combined active appearance and shape model is then denoted as shown in equation (2.8).

$$p = \begin{pmatrix} W_v O_v^T (v - \bar{v}) \\ O_a^T (a - \bar{a}) \end{pmatrix} \tag{2.8}$$

where W_v represent the compensation used for difference between shape and intensity based models.

2.2.2 Algorithms for Facial Aging Analysis

Over the past one decade, there has been a growing interest in understanding the effect of increasibg age on face recognition performance. A number of research efforts have been made to mitigate the effects of facial aging on face recognition performance. A detailed survey on different computational methods for facial age modeling has been presented in [38]. Generally, the recognition of age separated face images involves two components: (i) demographic estimation (i.e., age group, gender, and race estimation) and (ii) recognition of age-separated face images. Established approaches on each of these two components are described below.

Demographic estimation: The estimation of age, gender, and race is generally termed as demographic estimation and finds numerous applications such as identifying asylum seekers with missing documents. Due to significance of the problem, demographic estimation has been studied extensively in the literature, such as [39], [40], [41], [42], [43], [44], [45], [46], [47], [48], [49], [50]. Existing demographic estimation approaches can be grouped into three main categories:

(i) Landmarks-based approaches

(ii) Texture-based approaches

(iii) Appearance-based approaches

The landmarks-based approaches, such as [39], [40], use the distance ratios between facial landmarks to express the facial shapes of different age groups. However manual annotation in facial landmark detection limits the usability of such approaches in automatic demographic estimation systems.

The texture-based approaches, such as [45], [46], [49], [50], utilize facial texture features, e.g., Local Binary Patterns (LBP), Gabor, and Biologically Inspired Features (BIF). Though used in many demographic estimation approaches, high feature dimensionality makes them less efficient.

The appearance-based approaches, such as [41], [47], [48], adopt facial appearance to discriminate among face images belonging to different demographic groups.

Recognition of age-separated face images: Recognition of age-separated face images is still a challenging research problem and finds numerous applications such as finding missing persons. Recently a number of face recognition methods have been proposed for age estimation, such as [51], [52] and recognition of age-separated face images, such as [53], [54], [55] ,[56], [57], [58], [59], [60], [61], [62]. Existing methods for recognition of age-separated face images can broadly be classified into two categories:

(i) Generative methods

(ii) Discriminative methods.

The generative methods rely on modeling of facial aging process. Some of the representative works in this category include [57], [58]. In [57], a sequence of face images is stored in a time-order to model the age patterns. In [58], a view-invariant 3D age modeling technique has been proposed to compensate the age variations for recognition across time lapse. Though such methods are successful in discriminating age-separated face images, yet involve complex age simulation process.

On the other hand, discriminative methods use multiple feature representations to improve recognition performance. Some of the recent representative works in this category include [13], [14], [59], [58], [62]. In [59], a multi-feature discriminant analysis based on local facial features is presented to recognize age-separated face images. In [13], Bacteria foraging fusion scheme is used to optimize the weights assigned to different facial regions to maximize the recognition performance. A multi-view discriminative learning (MDL) approach based on three different local feature descriptors is presented in [62]. In a most recent approach [14] reported in the literature, human perception based fusion strategy has been proposed to recognize age separated face images. While existing approaches have achieved remarkable performance in recognizing age-separated face images, yet there are two main limitations: (i) there is lack of features evaluation with temporal variations of face images, and (ii) most of the existing methods rely on a single feature representation and hence lack the discriminative information. For example appearance based facial features exhibit poor performance in discriminating age-separated face images due to their dependence on raw pixel intensities. Local features,

though used in many age invariant face recognition methods, can be manipulated by certain age groups using artificial extrinsic facial variations (e.g., by facial make up) [16].

2.2.3 Facial Asymmetry based Face Recognition Algorithms

Bilateral facial asymmetry has long been studied as a critical factor in face recognition [63], [64], [65], [66], [67], evaluation of facial attractiveness [68], social status of an individual [69] and expressions [70]. In [63], facial asymmetry has been used for expression invariant human identification with encoding of facial asymmetry using PCA based subspace method. In [65], asymmetric left or right faces have been used for recognition, particularly in the situations when full faces are not available. Bilateral facial symmetry has been used in [66] to calculate symmetry scores on a 3D database to compare men and women sub groups along with most symmetric and least symmetric sub groups. Based on vertical facial symmetry, left half faces have been used for face recognition in [67], for a comparable accuracy to full face images. In [71], an optical flow based facial asymmetry measurement is presented to facilitate face reconstruction and recognition. However, the role of facial asymmetry in recognizing age-separated face images and extracting demographic informative features, such as age group, gender and race is yet to be determined, which is focus of the current research.

2.3 Face Images Databases

A number of facial databases have been collected to evaluate the performance of different face recognition algorithms. A number of variations occur in face images of acquired databases including pose, illumination, expressions, and temporal changes. Keeping in view such variations, we have selected three publically available, large and comprehensive facial aging databases, the FERET, MORPH, and FG-NET. The choice is motivated by the following two facts: (i) all these databases contain aging variations suitable to conduct research on recognition of age-separated face images, as suggested in current dissertation, and (ii) standard benchmarks studies are available to compare the performance of proposed methods on these databases. The important statistics of these databases are as follows.

2.3.1 The FERET Database

FERET [73] is one of the largest publically available databases, which adresses the aging problem of face images with following two important attributes.

(i) A large number of frontal face images (3540 face images of 1196 subjects).

(ii) Age-separated face images with small and large temporal variations, without exaggerated illumination and expression variations so that recognition problem of such images can be well studied. The database, along with other image sets, contains a gallery and four probe sets of frontal face images. The gallery set, termed as fa set, consists of 1196 facial images. The four probe sets are fb, fc, dup I and dup II. The fb set contains 1195 images, captured with an alternate expression compared to the corresponding gallery images. The fc set contains 194 images captured under different illumination conditions. The dup I set contains 722 images, acquired between 0 to 1031 days after the corresponding gallery images. Dup II is a subset of dup I and contains 234 images acquired at least one and half years after the corresponding gallery images. The average number of images per subject is 2.97 and 3.12 in dup I and dup II sets respectively with ages range from 10-70 years. Figure 2.3 shows example face images of a subject from FERET database with different facial variations including expression, illumination and temporal changes. Some important characteristics if FERET database are given in Table 2.1.

(a) (b) (c) (d) (e)

Figure 2.4: Fontal face images from FERET database (a) f_a (b) f_b (c) f_c (d) dup I (e) dup II set

Table 2.1: Important statistics of FERET database

	fa set	fb set	Dup I set	Dup II set
# Subjects	1196	1195	243	75
# Images	1196	1195	722	234
# Average images/subject	1	1	2.97	3.12
Temporal variations	N/A	N/A	Images were acquired later in a time with maximum temporal variation of 1031 days compared to corresponding gallery images	Images were acquired at least 1.5 years later than corresponding gallery images
Age range	10~60+			

2.3.2 The MORPH Database

MORPH [27] is another large facial aging database which contains 55134 face images of more than 13000 individual subjects and is extensively used in research on recognizing age-separated face images [59], [62]. The average number of images per subject is 4 in the age range of 16-77 years. The minimum and maximum temporal variations among images of an individual are 1 day and 1681 days respectively. In addition to temporal variations, the database also provides gender and race information of subjects. Some example age-separated face image pairs, with gender and race information, are shown in Figure 2.4.

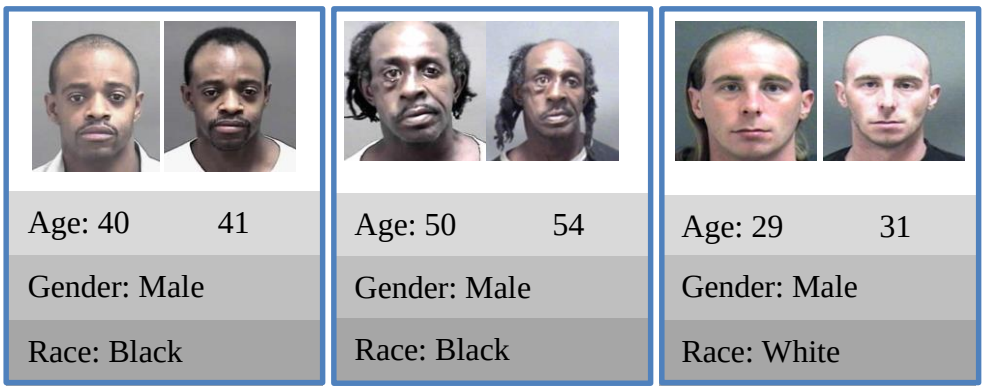


Figure 2.5: Example age-separated face image pairs with gender and race from MORPH database (age in years)

2.3.3 The FG-Net Database

FG-NET aging database contains 1002 age-separated face images of 82 subjects in the age range of 0-69 years. FG-NET database is extensively used in research studies focusing on recognition of age-separated face images such as images [59], [62]. Some age-separated face images from these databases are shown in Figure 2.5.

Age 14 28 50 54 8 12

Figure 2.6: Example age-separated face image pairs from FG-NET database (age in years)

2.4 Summary

This chapter presents an overview of face recognition specifics along with face normalization of a typical face recognition system. A brief but wide-ranging literature review of face recognition algorithms is also presented. We also describe the facial aging databases along with important characteristics of each database used in this dissertation.

Chapter 3

FEATURE REPRESENTATION AND EXTRACTION

The goal of feature representation is to encode different facial characteristics into facial feature vector descriptors. Feature extraction on the other hand focuses on feature combinations which project the original feature descriptors into a feature space with improved face class separation. Generally, feature extraction can be carried out either as a holistic representation or a local representation. In this chapter we have summarized the most relevant feature representation and feature extraction techniques employed in face recognition technology.

3.1 Feature Representation

The study presented in [15] has organized the ample facial feature representations employed in manual and automated face recognition into following three levels: (i) Level 1, (ii) Level 2, and (iii) Level 3 features. This classification of facial features is aimed at providing a clear understanding of feature extraction techniques. The three levels are described in the following subsections.

3.1.1 Level 1 Features

These features cover the holistic representation of the face. Level 1 features mainly focus at appearance based methods including eigenfaces [9] and fisherfaces [10]. Such features are able to differentiate general face occurrences like:

- A fat and elongated face.
- Male and female faces which possess certain predominant characteristics.
- Individual faces belonging to different origins.

Limitations: Level 1 features however, suffer from the fact that they are unable to classify an individual over a larger population. This fact has been explained in Figure 3.1.

Figure 3.1: Limitation of level 1 facial features. Level 1 features can discriminate between images with large differences but cannot differentiate similar faces of different subjects. Query face image (a) is matched with gallery images (b-d) using PCA with MahCos distance metric. Level 1 features well matched image (a) with image (c) of another subject with similarity to (a), than its true match (b), but differentiated query image (a) and gallery images (d) with large facial differences. (Images selected from MORPH database [27])

Besides their limitations, first level features are still useful for reducing the search space. These features can effectively be used in early stages of face recognition systems to improve the matching speed.

3.1.2 Level 2 Features

Level 2 features require more detailed observations while being used in face recognition tasks, compared to Level 1 features. As opposed to holistic approach, these features tend to derive local features from specific structures in the face which are the most relevant to recognition task. Examples of s such features include: (i) Gabor wavelets [73], (ii) Local Binary Patterns (LBP) [11], Scale Invariant Feature Transform (SIFT) feature descriptors [12], and Biologically Inspired Features (BIF) [74].

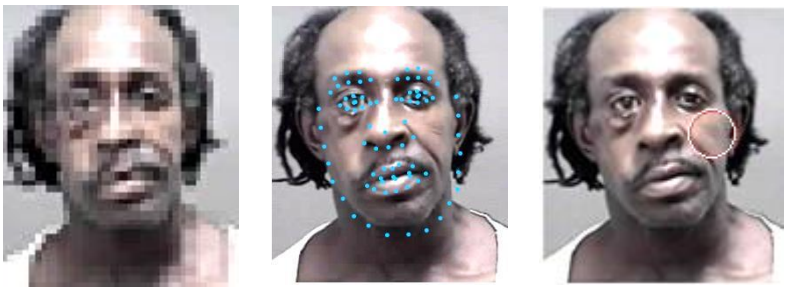
Limitations: Despite the fact that such features are generally discriminative in nature, yet under some conditions, they alone are not capable of recognizing faces e.g. in case of

face recognition of monozygotic twins with similar facial appearance and age invariance, where texture of face changes significantly with time lapse [15].

3.1.3 Level 3 Features

This class of features focuses on very small facial features, including scratches, wounds, and scars etc. Such information has been used for face recognition in [15]. These features have been used effectively in a challenging face recognition problem of monozygotic twins (identical twins). Due to high resemblance between identical twins, the presence of any scar or mark on the face may be an important cue towards recognition. Medical research has shown monozygotic twins may have same number of moles on their faces but their location is different. Level 3 features have been shown to be most robust in recognition tasks for such scenarios [75]. Figure 3.2 illustrate the above mentioned feature levels.

By comparison, it is observed that no optimal feature representation exist for face recognition task. Each feature level description has its own advantages and disadvantages. The above three facial feature extraction levels have been summarized in Table 3.1.



(a) (b) (c)

Figure 3.2: Examples of facial features levels (a) Low dimensional Level 1 features (b) Level 2 features describing shape and texture of face (c) Level 3 facial features representing scars, moles, tattoos and other facial irregularities

Table 3.1: Example features from three different facial feature levels to represent face images

Features	Examples
Level 1 features	Appearance-based methods e.g. PCA, LDA, etc.
Level 2 features	Local feature descriptors e.g. LBP, SIFT, etc.
Level 3 features	Very small facial features, such as scars, wounds, and moles etc.

3.2 Feature Extraction

This section gives a brief overview of the subspace algorithms used for feature extraction to perform subsequent face recognition. Subspace face recognition algorithms are the most explored methods in the holistic face representation [9], [10]. Such methods operate directly on an image representation and extract relative features. These methods reduce the dimensionality of the face images while retaining the statistical separation between different classes. These algorithms are known to as subspace methods because they project the images to lower dimensional space to perform recognition. Maximum possible information is retained while converting the images to lower dimensional space. The property of retaining maximum information is what prioritizes them over each other. There are three main steps in extorting feature vectors to recognize face images namely training, projection, and recognition.

In the training phase, the basis vectors of the subspace for each algorithm are calculated and saved. During projection, these basis vectors are loaded and then all the database images are projected onto these basis vectors which convert them to the lower dimensional subspace. These projected images are saved as templates to be used later in distance calculation in the recognition phase. The whole process is shown in Figure 3.3.

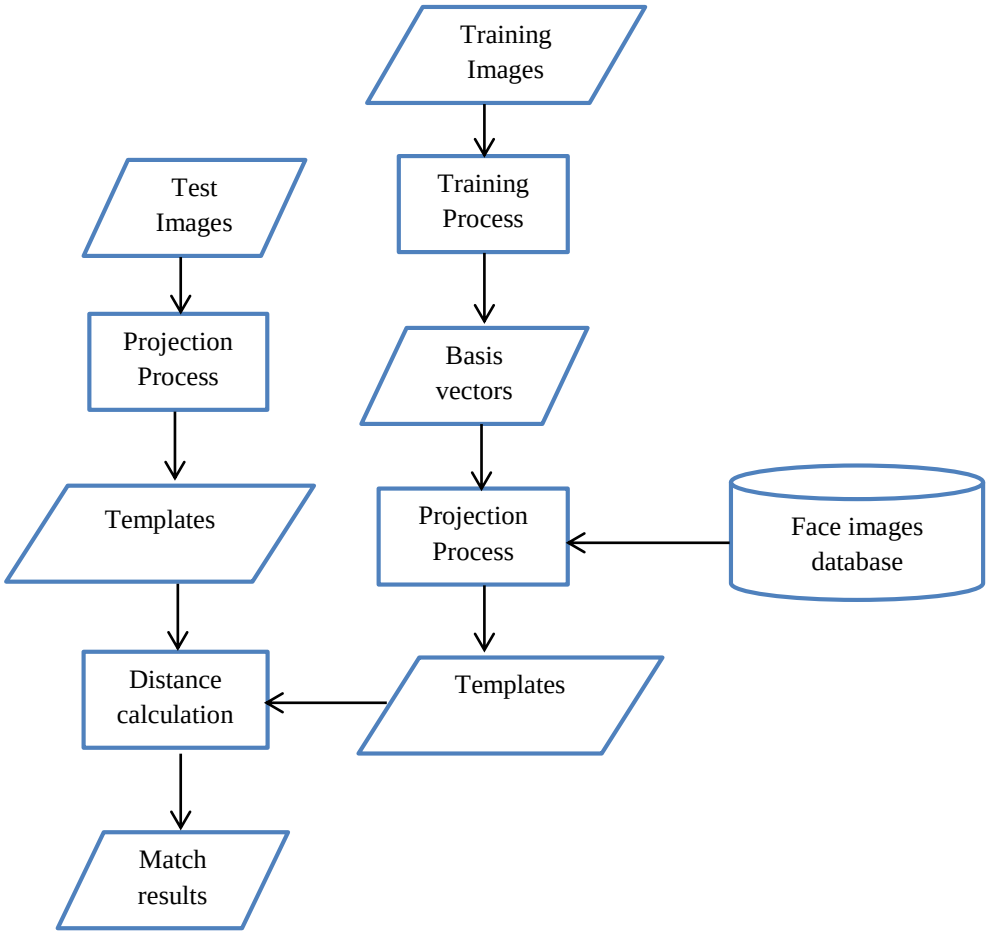


Figure 3.3: Face recognition process

A summarized description of the five well known subspace projection algorithms used to extract facial features is given in the following subsections.

3.2.1 Principal Component Analysis (PCA)

Principal Component Analysis (PCA) [9] is used to extract holistic facial features. Eigen faces are a set of eigen vectors that are broadly used in face recognition algorithms. In PCA a set of eigen vectors is derived, as described in the following steps.

(i) Consider N number of training images with $a \times b$ as size of each image. Suppose F be a matrix with each column containing mean-subtracted images. Then the covariance matrix C can be calculated as shown in equation (3.1).

$$C = FF^T \tag{3.1}$$

(ii) The eigenvector decomposition of C is given as shown in equation (3.2).

$$Cv_i = FF^T v_i = \lambda_i v_i \tag{3.2}$$

where v_i and λ_i are corresponding eigen vectors and eigen values, respectively.

But the size of covariance matrix C will be very large. On the other hand calculating $F^T F$ as covariance matrix, the eigen vector decomposition can be expressed as shown in equation (3.3).

$$F^T F u_i = \lambda_i u_i \tag{3.3}$$

Pre-multiplying both sides of equation (3.3) by F, we get equation (3.4).

$$FF^T F u_i = \lambda_i F u_i \tag{3.4}$$

Equation (3.4) implies that if u_i is an eigen vector of $F^T F$, then $v_i = F u_i$ is an eigen vector of C with manageable size.

(iii) Once the eigen vectors of C are calculated, the basis vectors D of size $a \times b$ can be calculated and normalized as shown in equation (3.5).

$$D = F \times Eigenvectors\ of\ C \tag{3.5}$$

(iv)The Eigenvectors corresponding to l largest eigen values are then selected out of total vectors of D and are termed as eigen faces, which in turn form a projection matrix R of the size $ab \times l$.

(v) Finally the templates, G of the size $l \times Q$ may be calculated as shown in equation (3.6).

$$G = R^T \times I \tag{3.6}$$

where I represents the Q number of image vectors to be projected in to R.

The major benefit of PCA is dimension reduction of a feature vector from d to d'. However PCA lacks the class specific information while performing subspace projection of underlying data.

3.2.2 Linear Discriminant Analysis (LDA)

PCA is effective only in reducing feature vector dimensionality and lacks the class specific information. Linear Discriminat Analysis (LDA) introduced in [10], is a face recognition subspace technique that seeks a subspace such that interclass separation is maximized while intra class separation is minimized, using fisher criterion defined in equation (3.7).

$$\phi = \underset{\phi'}{argmax} \frac{\phi^T S_B \phi}{\phi^T S_w \phi} \tag{3.7}$$

where S_B is between-class scatter matrix and S_w is within-class scatter matrix.

In order to learn an LDA based subspace projection, a training set of face images with n_s number of subjects is used with at least two images per subject. For a given i^{th} subject, μ_i is calculated as shown in equation (3.8).

$$\mu_i = \frac{1}{n_i} \sum_{j=1}^{n_i} x_i^j \tag{3.8}$$

Using equation (3.8), between-class scatter matrix S_B and within-class scatter matrix S_W are calculated as shown in equations (3.9) and (3.10), respectively.

$$S_B = \sum_{i=1}^{n_s} n_i \, (\mu_i - \mu)(\mu_i - \mu)^T \tag{3.9}$$

$$S_W = \sum_{i=1}^{n_s} \sum_{j=1}^{n_s} n_i (x_i^j - \mu_i)(x_i^j - \mu_i)^T \tag{3.10}$$

where μ represents the mean of all x_i^j feature vectors. Finally, in order to calculate ϕ, the generalized eigenvalue problem can be solved as shown in equation (3.11).

$$S_B \phi = \wedge S_W \phi \tag{3.11}$$

The benefit of LDA is dimension reduction of a feature vector from d to d' with class specific information being preserved. However LDA suffers from small sample size problem (SSS) [77] i.e. number of training samples is too small compared to feature vector dimensions.

3.2.3 Random Sampling Linear Discriminant Analysis (RSLDA)

Random Sampling Linear Discriminant Analysis (RDLSA) can be used to solve small sample size problem as suggested in [78]. RSLDA is employed to mitigate the effect of small sample size problem (SSS) by decomposing the feature space into more compact and solvable subsets like an ensemble classifier. LDA is then performed on reduced feature space with only a portion of original subjects used in each subspace.

3.2.4 Independent Component Analysis (ICA)

Independent component analysis (ICA) [79] is another subspace projection technique, widely used for feature extraction. ICA is a generalization of PCA, which is sensitive to higher order pixels relationships instead of only second order relationships. ICA is intimately related to blind source separation (BSS) problem, where goal is to decompose an observed signal into a linear combination of unknown independent signals. Let s and x represent the vectors of unknown source signals and observed mixtures, respectively, then the matrix model is represented as shown in equation (3.12).

$$x = As \tag{3.12}$$

where A is an invertible unknown mixing matrix. Based on the assumptions and observed mixtures, ICA tends to find mixing matrix A or separating matrix W as shown in equation (3.13).

$$u = W_x = WA_s \tag{3.13}$$

Equation (3.13) represents an estimation of independent source signals.

3.2.5 Laplacianfaces (LPP)

Laplacianfaces [80] is a subspace algorithm that projects face image data from high dimensions to lower dimensions while preserving the locality information of feature space. The face subspace is obtained by employing locality preserving projection after

performing the dimensionality reduction of face-space to compute an optimized embedding of the data points. A face image is mapped to a low dimensional face subspace using the locality preserving projection and is described by a combination of feature images known as laplacianfaces similar to the concept of Eigenfaces against PCA. The objective function of locality preserving projection is described as

$$\min \sum_{ij} (y_i - y_j)^2 S_{ij} \tag{3.14}$$

where y_i represents the one dimensional representation of an input image x belonging to the data matrix *A*, while *S* represents the adjacency matrix.

3.3 Summary

A comprehensive account of facial features representation and extraction used in face recognition algorithms is presented. Specifically, a classification of facial features representation has been presented to differentiate among plethora of presented approaches in literature. Major feature extraction schemes have also been discussed with advantages and limitations of each.

Chapter 4

DENSELY SAMPLED ASYMMETRIC FEATURES

The goal of this chapter is to find a specific representation of facial asymmetry that can highlight relevant information. This representation can be found by expressing a face image by a set of feature vectors where each feature vector summarizes the underlying content of a region by using local representation. These feature vectors can be projected into a new feature space, then the least relevant features can be removed to reduce the dimension of feature vector according to certain criterion.

In this chapter we give an overview of facial asymmetry etiology and propose a feature extraction scheme suitable to extract facial asymmetry based features, to be used in demographic estimation and recognition of age separated face images tasks.

4.1 Etiology of Facial Asymmetry

Facial asymmetry, which refers to non-correspondence in shape, size and arrangement of facial landmarks on both sides of the face, is common in humans, even in young healthy subjects [17]. Facial asymmetry may be classified as intrinsic and extrinsic asymmetry [81]. Intrinsic asymmetry corresponds to bilateral structural variations of face while extrinsic asymmetry is related to external facial variations, such as pose and illumination. Causes of intrinsic facial asymmetry in humans can generally be classified into four categories: (i) acquired, (ii) functional, (iii) congenital, and (iv) developmental [82]. As an intrinsic characteristic of face and its presence in general population, an evaluation of intrinsic facial asymmetry across temporal variations and its potential role towards recognizing age-separated face images is presented in the current study.

4.2 Facial Asymmetry based Densely Sampled Asymmetric Features

In order to extract asymmetric facial information, Densely Sampled Asymmetric Features (DSAF) are proposed. The proposed feature extraction scheme consists of three steps

with extraction of (i) difference half face image, (ii) facial asymmetries from difference half face image, and (iii) geometrical moment invariants to characterize facial asymmetries. An overview of the steps involved in extraction of proposed DSAF features from a preprocessed face image is shown in Figure 4.1. It can be seen that first the face images are preprocessed and a difference half face image is extracted for each face image. Then facial asymmetries are extracted which are characterized using geometrical moment invariants, resulting into DSAF features. The details of each step are given in the following subsections.

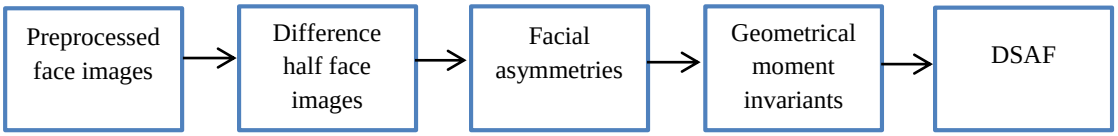


Figure 4.1: An overview of the steps involved in extraction of proposed DSAF

4.2.1 Extraction of Difference Half Face Images

To start with, each face image is preprocessed by upright rotation, alignment, illumination correction, and cropping. Each preprocessed face image is then divided into two parts; the left half face (LHF) and right half face (RHF). The mirror image of LHF is subtracted from RHF to extract a difference half face image as illustrated in Figure 4.2.

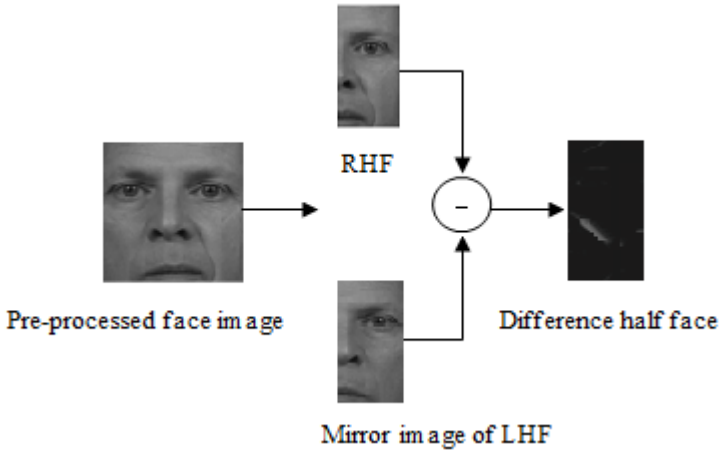


Figure 4.2: Extraction of difference half face image

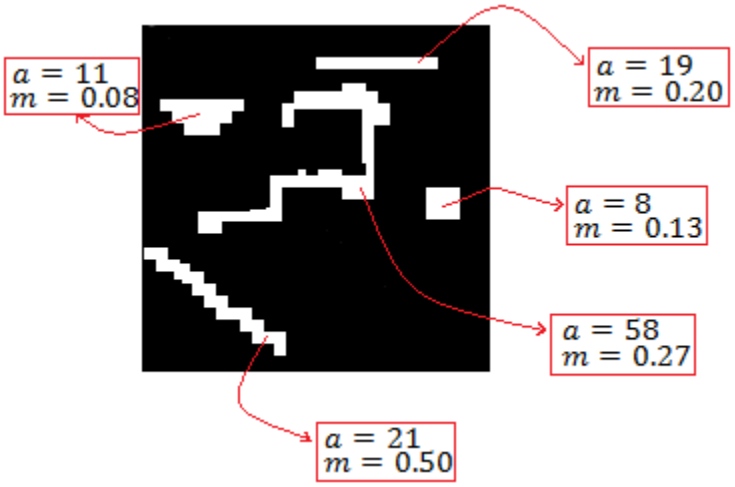
4.2.2 Extraction of Facial Asymmetries

As shown in Figure 4.2, each difference half face image, presents facial asymmetries (termed as asymmetries in the rest of text), which are mainly composed of iso-intensity pixels (pixels with same intensity). Such pixels can be straightforwardly grouped together and considered as unique recognition-specific asymmetries. From this point of view, it is possible to extract such asymmetries using attribute profiles (APs). APs as an extension of morphological profiles (MPs) [83], are used as an advanced mechanism for a multilevel analysis of image by the sequential application of morphological attribute filters (AFs). AFs are connected component (CC) transformations, used to model different kinds of the structural information. Recall that connected component is a region of iso-intensity connected pixels in an image. AFs can process an image by removing or preserving connected components according to a given criterion. They are flexible operators since they can transform an image according to many different attributes. An attribute is any measure computable on the regions (e.g. area and perimeter). AFs have two important criterions, the increasing and non-increasing criterion. If a criterion is satisfied for a connected region *K*, it will also be satisfied for all those regions that include *K*, it is called increasing criterion (e.g., area and length of bounding box of a region), otherwise it is called non-increasing criterion (e.g., moment of inertia and standard deviation of pixels intensities). If the criterion considered is increasing, the resulting transformation is increasing, (i.e., it is an *opening*).

In contrast, if the increasingness property is not fulfilled by the criterion, the transformation is called a *thinning*. Analogous considerations can be made for the dual transformation. If the criterion is increasing, the transformation is actually a *closing* otherwise it is a *thickening*.

A toy example of binary attribute thinning is shown in Figure 4.3. The filtered images show how the elongated objects can be isolated from other compact objects by using moment of inertia (m) attribute (Figure 4.3 (b)). By considering only the area attribute (a) it is possible to keep one of the two elongated objects (Figure 4.3(c)). The toy

example shows how complimentary information can be retained using different attributes.



(a)

(b)

(c)

Figure 4.3: Example of attribute thinning and opening (a) Original binary image showing the values of area (a) and moment of inertia (m) attributes for each region in the image (b) binary attribute thinning with $m \geq 0.20$ (c) binary attribute thinning with $a \geq 15$

The filtering operation implemented in APs is based on the evaluation of how a given attribute (attr) is computed for every connected component of an image I for a given threshold value λ. For a connected component (CC) of the image, if the attribute meets a

predefined condition (e.g. attr (CC) > λ), then the component is preserved, otherwise it is removed.

Given an ordered sequence of *n* number of thresholds, $\lambda_1, \lambda_2, \dots, \lambda_n$, an AP is obtained by applying a sequence of attribute thickening and thinning transformations to image I as shown in equation (4.1).

$$AP(I) = \{\emptyset^{u_{\lambda_L}}(I), \emptyset^{u_{\lambda_{L-1}}}(I), \dots, \emptyset^{u_{\lambda_1}}(I),\ I, \gamma^{u_{\lambda_1}}(I), \dots, \gamma^{u_{\lambda_{L-1}}}(I), \gamma^{u_{\lambda_L}}(I)\} \quad (4.1)$$

where $\emptyset$ and γ represent the thickening and thinning transformations, respectively, $u = \{u_\lambda : \lambda = 0, \dots, n\}$ represents the family of non-increasing criteria which also fulfills the property of increasingness, and L represents the level of corresponding transformation, with a total of $2L + 1$ levels. To extract asymmetries, four AFs are computed on each difference half face image by considering corresponding attributes of area (*a*), length of diagonal of bounding box (*d*), moment of inertia (*m*), and standard deviation of pixels intensities (*t*), resulting into four APs: area attribute profile (AP_a), length of diagonal of bounding box attribute profile (AP_d), moment of inertia attribute profile (AP_m) and standard deviation of pixels intensities attribute profile (AP_t).

It is observed that both the conventional MPs and attribute profiles perform multi-scale image analysis, since any particular criterion, driven by the increasing values of the scalar quantity λ, progressively erases from the image larger structures. Moreover, attribute opening profiles provide the same capabilities as for openings by reconstruction in processing the image but with greater flexibility in terms of filtering criterion. For example, if we consider a compact structuring element (e.g., square or disk-shaped), the structures are removed from the scene if the structuring element does not fit in them. Thus, the image is processed according to the smallest size of the regions. On the other hand, if we consider the attribute of the area of the regions present in an image, then the structures are filtered according to a measure different from the smallest size of the regions. Similarly, if we take into account the length of the diagonal of bounding box of a region, then a different measure of the size of the objects is provided. Thus, by selecting

different type of attributes, different characterizations of the scale of the structures are generated.

Thus, to perform multilevel image analysis, different values of scalar λ are used. To create AP_a following threshold values of λ are used: 2, 6, 15, and 30. For AP_d, the following threshold values of λ are selected: 2, 4, 6, and 12. To create AP_m the difference half face image is filtered by progressively suppressing those regions with corresponding attribute values smaller than the following increasing thresholds: 0.3, 0.5, 0.8, and 0.9. AP_t models the homogeneity of the gray levels of the pixels in the image. The profile is built according to the following reference values of the standard deviation: 4, 6, 8, and 10.

To consider the effect of all computed APs simultaneously, combined attribute profile AP_c is defined as shown in equation (4.2).

$$AP_c = \{AP_a, AP_d, AP_m, AP_t\} \tag{4.2}$$

Figure 4.4 shows the application of AP_c to a sample input difference half face image. The output difference half face image shows the extracted asymmetries with iso-intensity pixels (white pixels).

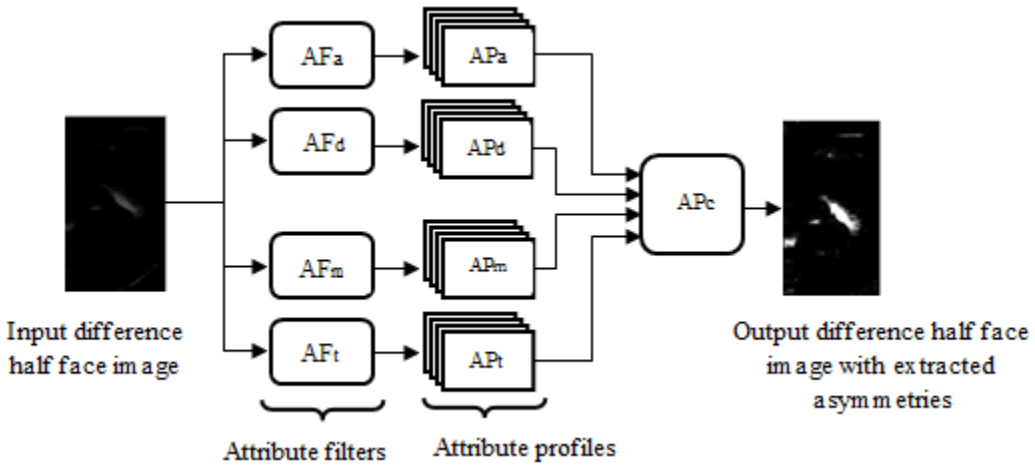


Figure 4.4: Proposed approach to extract facial asymmetries using attribute profiles (APs)

4.2.3 Extraction of Geometrical Moment Invariants

Each extracted asymmetry can be geometrically characterized by computing its geometrical moment invariants, introduced by Hu [84] and termed as Hu moment invariants in the rest of text. The Hu moments can be used to describe, characterize, and quantify the shape of an object in an image, with reasonable feature vector length. The geometrical moment technique is used to derive such invariant moments as described below.

Consider 2-D continuous image function $f(x, y)$, the moments, T_{pq} of the order $(p + q)$ are given as shown in equation (4.3).

$$T_{pq} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x^p y^q f(x, y) dx dy \tag{4.3}$$

For a discrete grey scale image with pixel intensities $P(x, y)$, the image moments can be defined as shown in equation (4.4)

$$T_{mn} = \sum_m \sum_n x^m y^n P(x, y) \tag{4.4}$$

The central moments for grey scale image are then defined as shown in equation (4.5).

$$\lambda_{mn} = \sum_m \sum_n (x - x_c)^m (y - y_c)^n P(x, y) \tag{4.5}$$

where x_c and y_c are centroid components, defined as

$$x_c = \frac{T_{10}}{T_{00}} \quad and \quad y_c = \frac{T_{01}}{T_{00}}$$

The normalized central moments are defined as given in equation (4.6).

$$\mu_{mn} = \frac{\lambda_{mn}}{\lambda_{00}^{\gamma}} \tag{4.6}$$

where

$$\gamma = \frac{(m + n + 2)}{2}, \qquad p + q = 2,3, \ldots$$

Based on normalized central moments given in equation (4.6), Hu defined following seven moment invariants as given in equations (4.7-4.13).

$$T_1 = \mu_{20} + \mu_{02} \tag{4.7}$$

$$T_2 = (\mu_{20} + \mu_{02})^2 + 4\mu_{11}^2 \tag{4.8}$$

$$T_3 = (\mu_{30} - 3\mu_{12})^2 + (3\mu_{21}^2 - \mu_{03})^2 \tag{4.9}$$

$$T_4 = (\mu_{30} + \mu_{12})^2 + (3\mu_{21} + \mu_{03})^2 \tag{4.10}$$

$$T_5 = (\mu_{30} - 3\mu_{12})(\mu_{30} + \mu_{12})[(\mu_{30} + \mu_{12})^2 - 3(\mu_{21} + \mu_{03})^2] + (3\mu_{21} - \mu_{03})(\mu_{21} + \mu_{03})[3(\mu_{30} + \mu_{12})^2 - (\mu_{21} + \mu_{03})^2 \tag{4.11}$$

$$T_6 = (\mu_{20} - \mu_{02})[(\mu_{30} + \mu_{12})^2 - (\mu_{21} + \mu_{03})^2] + 4\mu_{11}(\mu_{30} + \mu_{12})(\mu_{21} + \mu_{03}) \tag{4.12}$$

$$T_7 = (3\mu_{21} - \mu_{03})(\mu_{30} + \mu_{12})[(\mu_{30} + \mu_{12})^2 - 3(\mu_{21} + \mu_{03})^2] - (\mu_{30} - 3\mu_{12})(\mu_{21} + \mu_{03})[3(\mu_{30} + \mu_{12})^2 - (\mu_{12} + \mu_{03})^2] \tag{4.13}$$

These seven moments are invariant to rotation, translation, and scale changes.

In order to extract Hu moment invariants, each difference half face image with extracted asymmetries is divided into a set of overlapping patches and geometrical moment invariants are computed for each patch, resulting into Densely Sampled Asymmetric Features (DSAF), as described below. Consider a difference half face image with extracted asymmetries, of size $h \times w$, with patch size $p \times q$ and overlapping radius *r*, then the number of horizontal (X) and vertical (Y) patches obtained are given in equations (4.14) and (4.15), respectively.

$$X = \frac{h - r}{p - r} \tag{4.14}$$

$$Y = \frac{w - r}{q - r} \tag{4.15}$$

For each patch, a d-dimensional Hu moment invariant feature vector is extracted, resulting into a $d \times X \times Y$-dimensional DSAF feature vector for a given difference half face image with extracted asymmetries as illustrated in Figure 4.5.

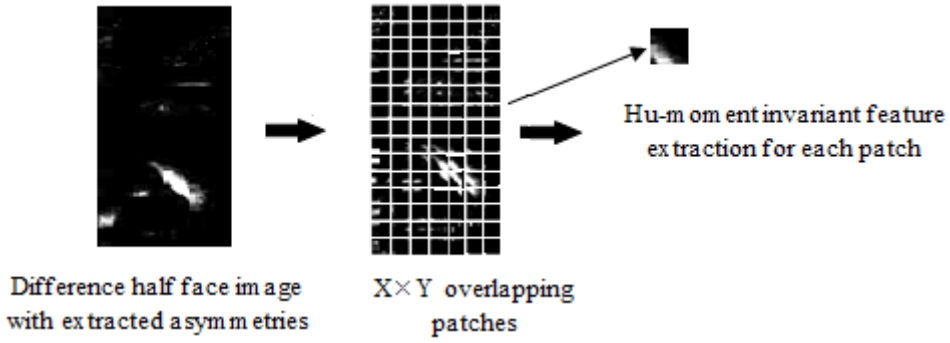


Figure 4.5: Hu moment invariant feature extraction from each patch

The resulting DSAF features can be used to match two similar asymmetries in two different images of the same subject in face recognition task.

4.3 Comparison of DSAF with Existing Methods on Extraction of Facial Asymmetry

To check the performance of proposed DSAF features in recognizing face images, we have compared DSAF features with PCA based representation of facial asymmetry presented in [63]. We perform face recognition experiments on FERET database, by selecting fa set as gallery set, while dup I and dup II as probe sets. In case of DSAF, a 735-dimensional feature vector is extracted for each difference half face image of size 128x64 with extracted asymmetries. Similarly, in case of PCA, a 60-dimensional feature vector is extracted for each 128x64 difference half face image as suggested in [63].

In case of DSAF features, we get rank-1 identification accuracies of 56.92 % and 36.75% on dup I and dup II sets, respectively. In case of PCA feature vectors, we achieve rank-1 identification accuracies of 41.55% and 19.23% on dup I and dup II sets, respectively.
The recognition accuracies suggest the superior performance of proposed DSAF features in recognizing face images compared to existing method presented in [63].

4.4 Summary

An account of etiology of facial asymmetry has been presented along with newly proposed Densely Sampled Asymmetric Facial Features (DSAF). We introduced DSAF as asymmetric facial features. By densely sampling the difference half face images, sufficient discriminatory information is available to classify face images. In the following chapters, we will analyze the performance of proposed feature descriptors in demographic estimation and recognition of age-separated face images in particular scenario of recognizing age-separated face images.

Chapter 5

RECOGNITION OF AGE-SEPARATED FACE IMAGES

In this chapter we present a rigorous analysis of measurement and evaluation of facial asymmetry with small and large temporal variations. The objective of this measurement and valuation is to analyze the impact of aging on facial asymmetry. The effectiveness of proposed DSAF features compared to some existing feature descriptors is assessed in typical scenario to recognize age-separated face images. In particular, we discuss the ability of Support Vector Machine (SVM) classifier to recognize age-separated face images in matching-score space (MSS).

5.1 Introduction

Human face contains a number of recognition-specific clues including age, gender, race and expressions. Over the past five decades, several face recognition algorithms have been proposed in the literature. Currently, face recognition of age-separated face images is receiving continuous attention of research community. The major challenge in recognition of age-separated face images is to extract such discriminative information, which can reduce matching scores gap between gallery and corresponding probe images. Facial temporal variations are highly complex in nature due to number of factors including facial asymmetry. Although facial asymmetry has been used in previous face recognition studies [65], [66], [67], [71], yet these studies lack the evaluation of facial asymmetry across temporal variations and its discriminative role in recognizing age-separated face images. Motivated by this fact, we aim at: (1) evaluation of facial asymmetry with small and large temporal variations; and (2) combining asymmetric facial features along with holistic (PCA) [9] and local feature descriptors (LBP) [11] to exploit the discriminative information for recognition of age-separated face images.

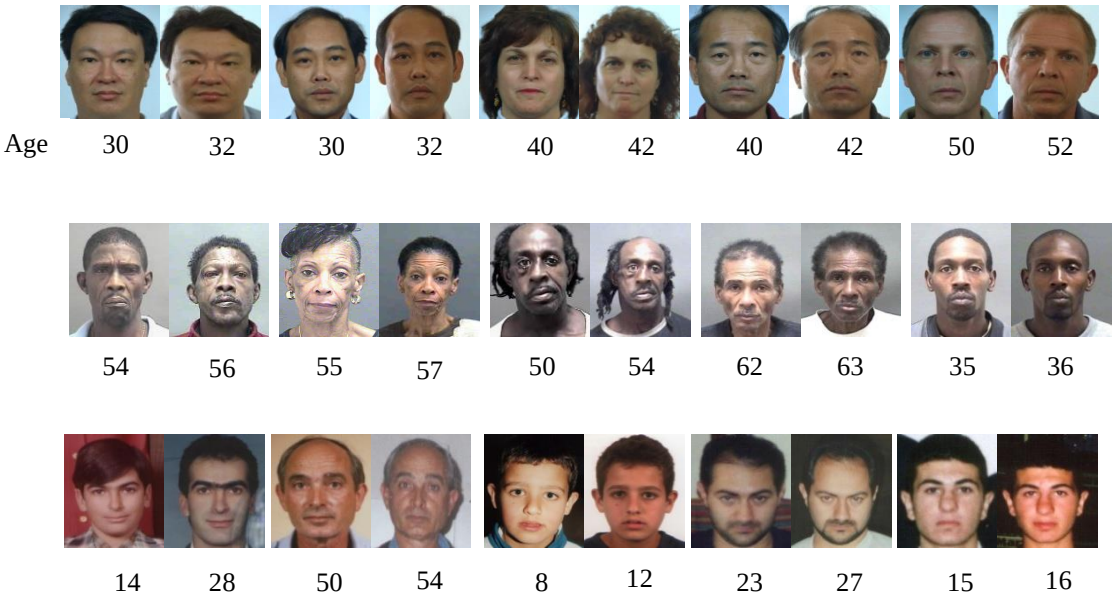


Figure 5.1: Example age-separated face image pairs showing intra-subject temporal variations: top row FERET database, middle row MORPH database, and bottom row FG-NET database

5.2 Facial Asymmetry Measurement and Evaluation across Temporal Variations

In this section, a comprehensive account of facial asymmetry measurement and evaluation is presented for small and large temporal variations on three widely used FERET, MORPH, and FG-NET databases, with necessary face pre-processing and landmarks detection given in the following subsections. Some age-separated face images from these databases are shown in Figure 5.1.

5.2.1 Face Pre-Processing and Landmarks Detection

Pre-processing: To compensate for different appearance variations, face images are pre-processed as described in following steps.

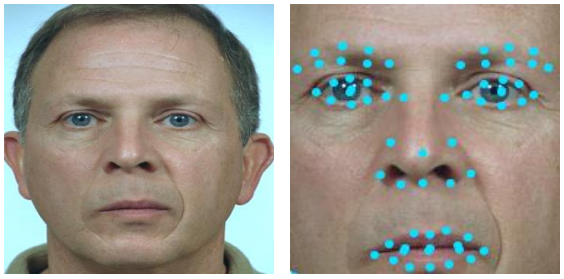
(i) Each face image is rotated such that it is vertically upright.

(ii) The face images are aligned based on eyes coordinates such that inter-pupillary distance is same for all images.

(iii) Illumination variations are corrected by using histogram equalization.

(iv) Each face image is tightly cropped (based on eye coordinates) as suggested in [62] into 128×128 pixels size, including middle of forehead to just below the mouth and from outer corner of one eye brow to the other, to remove unwanted background and hair regions.

Landmarks detection: Sixty-four facial landmarks are detected on each face image by following two methods: (i) manual annotation, and (ii) automatic landmarks detection, using Face++ [25] (a publically available state-of-the-art facial landmarks detection tool). To check the validity of both methods, we calculate average error in detecting facial landmarks, which results into 1 pixel per image. Keeping in view this minimal error, the automatic method for facial land mark detection is selected. A sample original face image and its cropped version with detected landmarks using Face++ is shown in Figure 5.2.



(a) (b)

Figure 5.2: (a) A sample original face image (b) cropped face image showing detected landmarks using Face++

5.2.2 Measurement and Evaluation of Facial Asymmetry

To analyze the relationship between facial asymmetry and increasing age, we are interested to measure and evaluate facial asymmetry across small and large temporal variations. For this purpose, face image sets with small and large temporal variations from FERET, MORPH, and FG-NET databases are selected. In case of FERET database, 75 face images are selected, each from fa, dup I (small temporal variations) and dup II (large temporal variations) sets, such that the face images of subject present in fa set are also present in dup I and dup II sets. In case of MORPH database, three subsets

consisting of 10000 face images each are selected, representing the youngest, older (small temporal variations) and oldest (large temporal variations) face images of selected subjects. Similarly for FG-NET database, we select three subsets consisting of 82 face images each, representing the youngest, older and oldest face images of selected subjects. Facial asymmetry is then measured and evaluated as described in following steps.

(i) A facial midline y_1y_2 perpendicular to x_1x_2 (joining eye coordinates) is determined for each face image, through an iterative process such that difference between left and right half face is minimum, as shown in Figure 5.3 (a).

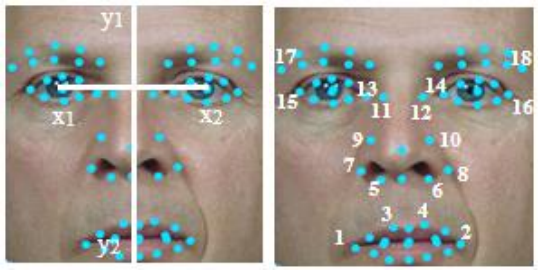
(ii) Nine pairs of bilateral facial landmarks (shown in Figure 5.3 (b), and illustrated in Table 5.1) are selected from facial regions which are least affected by facial expressions.

(iii) Facial asymmetry is then measured in terms of inter-pixel linear distances d_l and d_r of selected landmarks, reference to the facial midline y_1y_2. The difference D between d_l and d_r is given as shown in equation (5.1).

$$D = d_l - d_r \tag{5.1}$$

In equation (5.1), D gives asymmetry for respective bilateral landmarks. If D is positive then a given landmark will be skewed towards left and vice versa.

(iv) Absolute values of D are used to calculate mean and standard deviation (SD), which are recorded as bar graphs shown in Figure 5.4, for the selected subsets of FERET, MORPH and FG-NET databases.



(a) (b)

Figure 5.3: (a) Facial midline y_1y_2 perpendicular to line joining eye coordinates x_1x_2 (b) selected bilateral facial landmarks

Table 5.1: Illustration of bilateral facial landmarks [17]

Landmark	Abbreviation	#	Explanation
Cheilion	Chel	1, 2	The outer corner of the mouth, where the outer edges of the upper and lower vermilions meet.
Crista Philtre	Crist	3, 4	The point on the crest of the philtrum, the vertical groove in the median portion of the upper lip, just above the vermilion border.
Sub Alare	Sub-alar	5, 6	The point on the lower margin of the base of the nasal ala, where the ala disappears into the upper lip skin.
Alare	Alar	7, 8	Most lateral point on alar contour.
Lateral cartilage	Lat-Cart	9, 10	The base of upper lateral cartilage.
Maxillofrontale	Max	11, 12	The anterior lacrimal crest of the maxilla at the fronto-maxillary suture.
Endocanthion	End	13, 14	The inner corner of the eye fissure, where the eyelids meet.
Exocanthion	Exo	15, 16	The outer corner of the eye fissure where, the eyelids meet.
Frontotemporale	Fronto	17, 18	The most medial point on the temporal crest of the frontal bone.

In case of FFRET database, mean and standard deviation of fa set is compared with that of dup I and dup II sets, while mean and standard deviation of the youngest face images set is compared with that of older and oldest face images sets in case of both the MORPH and FG-NET subsets.

By comparison of thee results of facial asymmetry variation with small and large temporal changes, it is observed that:

- Facial asymmetry is more evident for the landmarks located on lower part of face than those for middle and upper parts of face.
- Facial asymmetry increases with age for all the selected image subsets. A slight increase in facial asymmetry is observed in lower part of face including sub-alare to crista philtre for small temporal variations. On the other hand, in case of large temporal variations we see a comparatively increased facial asymmetry, specifically for lower parts of the face including alare, sub-alare, crista philtre and cheilion on all selected subsets of face images. Keeping in view the earlier

related studies, our results are largely in line with those described in [82]. The evaluation results give us a clue that facial asymmetry is an intrinsic property of face and a strong indicator of age and hence can be used to recognize age-separated face images.

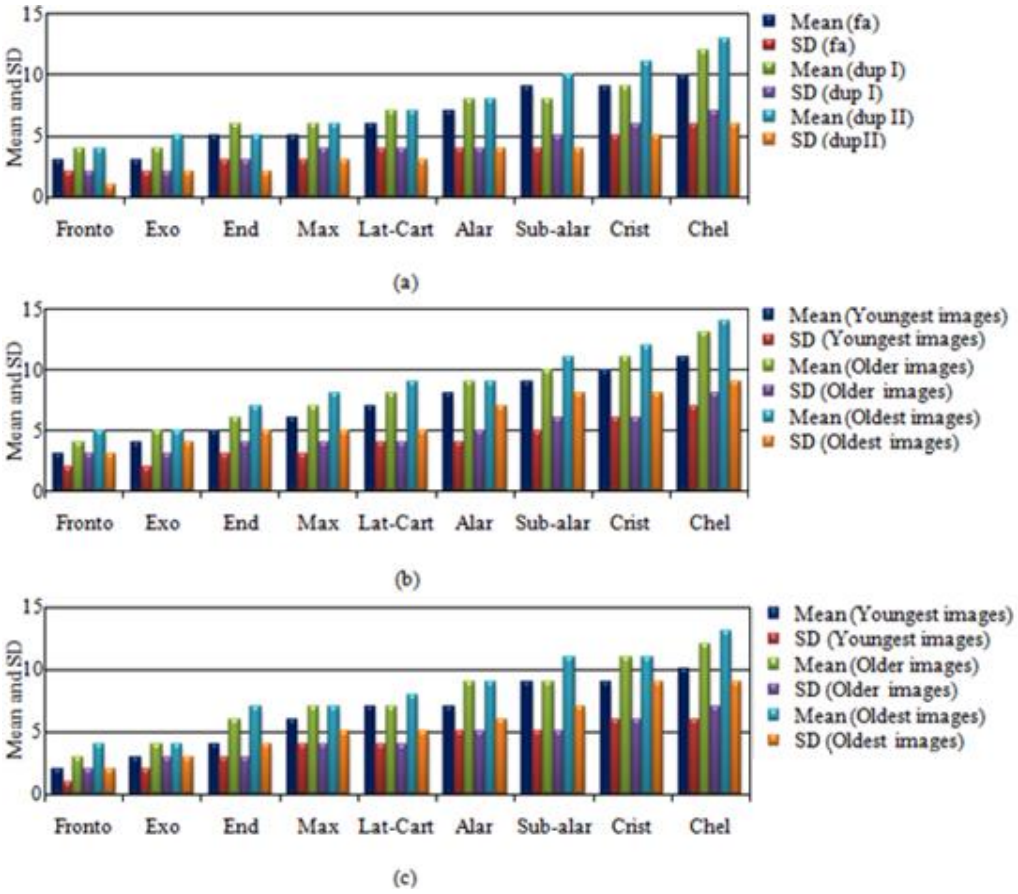


Figure 5.4: Facial asymmetry variation with temporal changes for (a) FERET, (b) MORPH, and (c) FG-NET subsets

5.3 Proposed Face Recognition Methodology

After assessment of facial asymmetry, we aim at using asymmetric facial features towards recognition of age-separated face images. The corresponding feature extraction and proposed face recognition algorithm is presented in this section.

5.3.1 Feature Representation

For recognition purpose, each face image is represented by extracting three feature descriptors including: (i) PCA based holistic features; (ii) LBP based local features; and (iii) facial asymmetry based Densely Sampled Asymmetric Features (DSAF).

PCA based holistic features: Principal Component Analysis (PCA) [9] is used to extract holistic facial features. Eigen faces are a set of eigen vectors, widely used in face recognition algorithms. In PCA, a set of eigen vectors is derived as follows.

Consider *N* number of training images with $a \times b$ as size of each image. Suppose F be a matrix with each column containing mean-subtracted face images. Then, as suggested by original study, the size of covariance matrix C can be reduced by calculating $F^T F$ instead of FF^T. Multiplication of F with the eigen vectors of $F^T F$ results into eigen faces, which are termed as basis vectors and form a projection matrix. In projection phase the required numbers of image vectors V are projected onto this projection matrix to achieve the templates.

LBP based local features: Local Binary Patterns (LBP) introduced in [11], can encode shape and texture of an underlying face image. These local facial features are robust to geometric and illumination variations. Uniform circular local binary patterns are extracted by dividing face image into square regions. An LBP histogram is then computed independently for each region. All the resulting histograms are then concatenated together to form a spatially enhanced histogram. A histogram of a labeled image $f(x,y)$ can be given as shown in equation (5.2).

$$H_i = \sum_{x,y} J\{f(x,y) = i\}, \qquad i = 0, \dots, l_1 - 1 \tag{5.2}$$

where $J(Z) = \begin{cases} 1, & Z \text{ is } true \\ 0, & Z \text{ is } false \end{cases}$

In order to retain the spatial information, the image is divided into regions, $R_0, R_1, R_2, \ldots, R_l - 1$ which results into a spatially enhanced histogram as shown in equation (5.3).

$$H_{i,j} = \sum_{x,y} J\{f(x,y) = i\}\, J\{(x,y) \in R_j\} \tag{5.3}$$

where $i = 0, 1, 2, \ldots, l_1 - 1$ and $j = 0, 1, 2, \ldots, l_2 - 1$ represent labels produced by LBP operator. The histogram in equation (5.2) contains information about corners, edges, and flat areas, while histogram in equation (5.3) results into a global description of face image.

Facial asymmetry based DSAF features: In order to use facial asymmetry as a feature descriptor, each original half face image is used to extract difference half face image and Densely Sampled Asymmetric Facial features are then extracted for each difference half face image as introduced in Chapter 4.

5.3.2 Proposed Face Recognition Algorithm

Figure 5.6 illustrates the block diagram of proposed face recognition algorithm, which is described in the following steps.

(i) The training images are partitioned into two datasets, dataset 1 and dataset 2, such that both datasets contains images of same subjects acquired under different conditions.

(ii) Each face image is pre-processed as described in section 5.2.

(iii) For each pre-processed face image, three different feature vectors are extracted as described blow.

(1) A 100-dimensional PCA based feature vector is extracted for each 128×128 face image.

(2) LBP based feature vector is extracted for each 128×128 face image by dividing it into 15×15 blocks with a patch size of 16×16 and overlapping radius of 8 pixels. We extract a 256-dimensioanl LBP feature for each patch resulting into a 15×15×256=57600-dimensional LBP feature vector for each face image.

(3) Each difference half face image with extracted asymmetries is divided into four regions, region 1, region 2, region 3, and region 4, of size 32×64 each. Each region is further divided into 7×3 blocks with a patch size of 16×16 and overlapping radius of 8 pixels. A 7-dimensional Hu moment invariant feature vector is extracted for each patch resulting into 3×7×7 =147-dimensional DSAF feature vector for each region as illustrated in Figure 5.5.

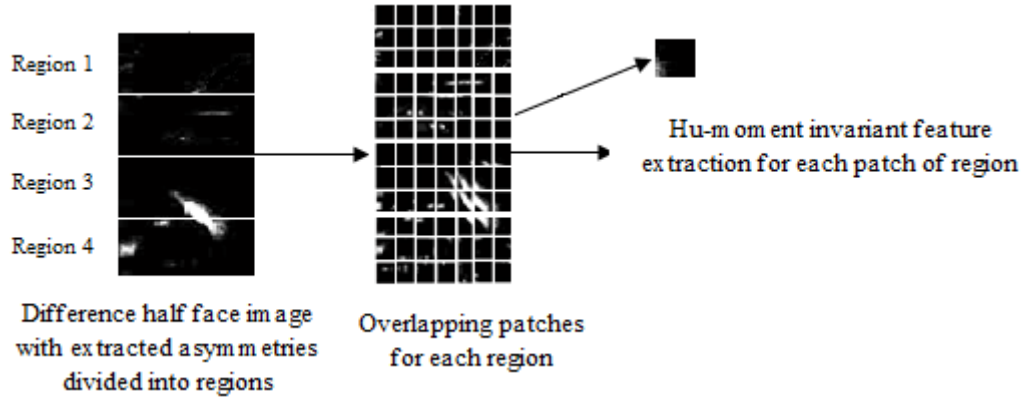


Figure 5.5: Extraction of DSAF features for each region of difference half face with extracted asymmetries

(iv) Once the feature vectors are extracted, the next step is to calculate matching scores using PCA, LBP, and DSAF features vectors. In case of PCA and LBP, respective feature vectors of images in dataset 1 are compared with those of dataset 2, to calculate matching scores. Genuine matching scores are obtained by comparing the feature vectors of same subjects while imposter matching scores are obtained by comparing the feature vectors of different subjects present in dataset 1 and dataset 2. In case of DSAF, a region-wise feature vector comparison of images in dataset 1 and dataset 2 is made to calculate genuine and imposter matching scores (e.g., feature vector of region 1 of an image in dataset 1 is compared with feature vector of same region of the same image present in dataset 2). Different weights are assigned to scores pertaining to each region (region 1, region 2, region 3, and region 4). The weights are calculated using the individual recognition accuracy achieved by each region R as shown in equation (5.4).

$$Weights(R) = \frac{Individual\ recognition\ accuracy\ of\ region\ R}{\sum Individual\ recognition\ accuracy\ of\ all\ regions} \quad (5.4)$$

Once the weights are assigned, the matching scores of all four regions are fused together to get combined DSAF matching scores S_w, using weighted sum rule, as shown in equation (5.5).

$$S_w = \sum_{i=1}^{4} w_i s_i \quad (5.5)$$

where s_i and w_i represent number of the scores and weights pertaining to each of four region, respectively. The resulting matching scores are termed as weighted DSAF matching scores. For A number of subjects in dataset 1 and B number of subjects in dataset 2, we get a matching-score matrix M_j of size $A \times B$, each for PCA, LBP, and DSAF feature vectors. Since A and B have same number of subjects in training, M_j is a diagonal matrix as shown in equation (5.6).

$$M_j = \begin{bmatrix} m_{11} & m_{12} & m_{13} & .... & m_{1A} \\ m_{21} & m_{22} & m_{23} & .... & m_{2A} \\ . & . & . & .... & . \\ . & . & . & .... & . \\ m_{B1} & m_{B2} & m_{B3} & .... & m_{BA} \end{bmatrix} \quad (5.6)$$

where each $m_{row,\ column}$ represents a matching score, with diagonal entries as genuine and non-diagonal entries as imposter scores, respectively. We get three matching score matrices M_1, M_2, and M_3 each for PCA, LBP, and weighted DSAF matching scores, respectively. Each matching-score matrix is normalized row-wise on a scale of 0 to 1 using simple *min-max* rule [31], to get a normalized matching-score matrix M'_{jrow} as shown in equation (5.7).

$$M'_{row} = \frac{M_{jrow} - \min(M_{jrow})}{\max\left(M_{jrow} - \min(M_{jrow})\right) - \min(M_{jrow} - \min\left(M_{jrow}\right))} \quad (5.7)$$

Where $\max(M_{jrow})$ and $\min(M_{jrow})$ denote the maximum and minimum values in a particular row of matching-scores matrix, respectively.

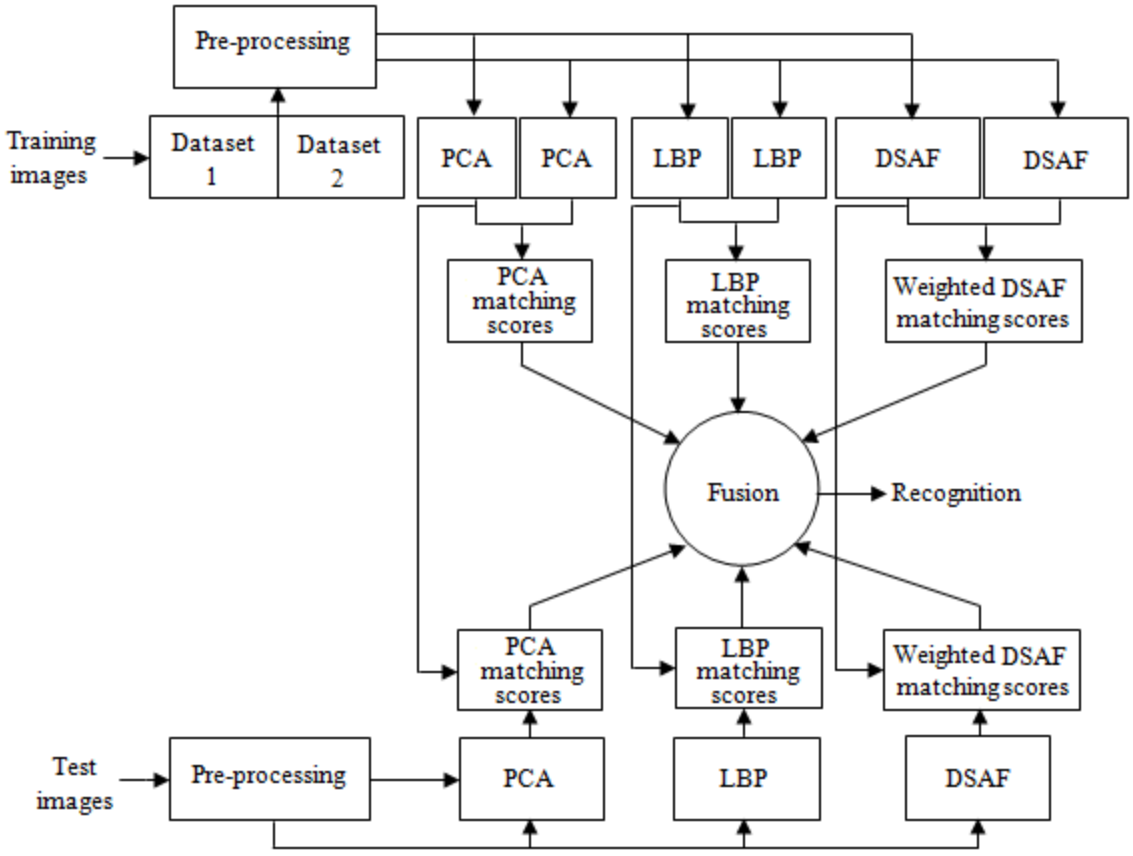


Figure 5.6: Block diagram of proposed face recognition algorithm

(v) The normalized matching-scores matrices M_1', M_2', and M_3', each for PCA, LBP and weighted DSAF matching scores are then fused to get a 3-dimensional matching-score space (MSS) by applying information combination approach [31]. In information combination approach, the matching scores from corresponding matrices form a score vector to represent a genuine or imposter class. Thus the vector (m_1', m_2', m_3') is a score vector, where m_1', m_2', and m_3', represent the matching scores of corresponding algorithms. An overview of the proposed matching-scores based fusion in MSS is shown in Figure 5.7.

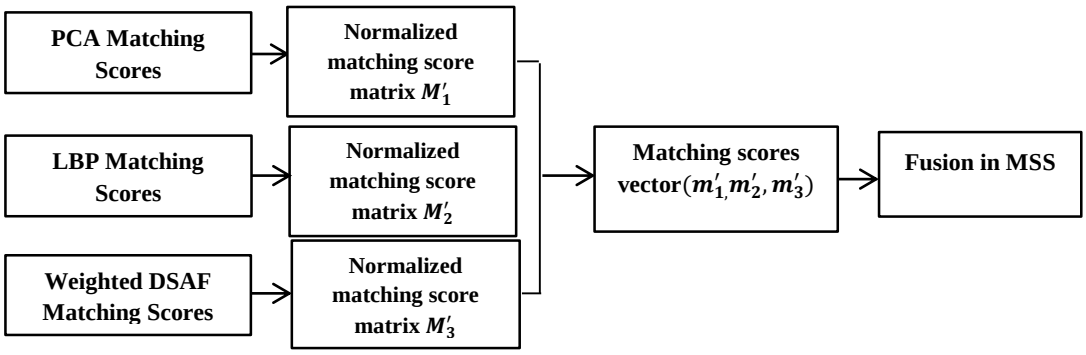


Figure 5.7: Block diagram of the matching-scores based fusion in MSS

(vi) Once the matching scores are fused in 3-dimensional MSS, they are classified either as genuine or imposter matching scores using Support Vector Machine (SVM) [85] as a binary classifier, for face verification task.

(vii) For face identification task, PCA, LBP, and weighted DSAF matching scores calculated in step (iv) are fused using simple sum rule to get a matching-score matrix M_k as shown in equation (5.8).

$$M_k = \sum_{k=1}^{n} M_k' \tag{5.8}$$

where n represents number of individual algorithms. The matching-scores matrix in equation (5.8) is normalized to calculate normalized matching-score matrix M_k', as shown in equation (5.9).

$$M_k' = \frac{M_k - \min(M_k)}{\max(M_k - \min(M_k)) - \min(M_k - \min(M_k))} \tag{5.9}$$

These normalized matching scores are used to determine recognition rates for face identification task.

(viii) In testing stage, we calculate the same set of feature vectors for a given probe set of face images, and compare with feature vectors of face images present in dataset 1 to calculate genuine and imposter matching scores. The matching scores are then used to perform face verification and face identification experiments.

5.4 Experiments and Results

The efficacy of proposed method is evaluated by conducting face verification and face identification experiments on FERET, MORPH and FG-NET databases with corresponding accuracies reported in terms of false acceptance rate (FAR) and identification rates, respectively. We plot the receiver operating characteristic (ROC) curves and cumulative match characteristic (CMC) curves, which are fundamental tools for experimental tests evaluation.

In a ROC curve, the false acceptance rate is plotted in function of the verification rate for different cut-off points. The ROC measures the quality of a 1:1 verification matcher. Each point on the ROC curve represents a FAR/verification pair corresponding to a particular decision threshold. The area under the ROC curve is a measure of how well a parameter can distinguish between two classes i.e., genuine and imposter.

CMC curve, on the other hand, is used as a measure of 1:m identification system performance. It judges the ranking capabilities of an identification system. The experimental setup and results on FERET, MORPH, and FG-NET databases are given below.

5.4.1 Experiments on FERET Database

(i) For face verification experiments, 501 pairs of frontal face images from fa and fb sets are used in training, resulting into 501 genuine and 250500 imposter matching scores each for PCA, LBP, and weighted DSAF algorithms. Dup I (722 images) and dup II (234 images) are used as probe sets for face verification against small and large temporal variations, respectively. A sample scatter plot showing genuine and imposter scores in proposed MSS is shown in Figure 5.8, for FERET training dataset. From scatter plot, it can be seen that:

- It is possible to combine matching scores from different algorithms in proposed MSS to represent genuine and imposter scores.
- Once combined, these scores are reasonably seperated in MSS and are classified by using SVM as a classifier with radial basis function (RBF) kernel. A publically

available libsvm [85] is used for implimentation of SVM in all face verification experiments.

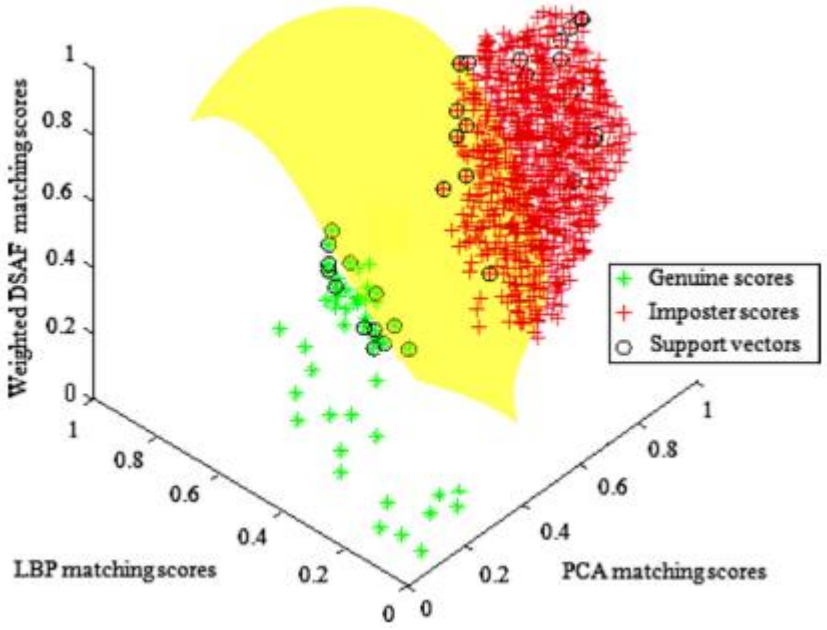


Figure 5.8: A sample scatter plot showing SVM based classification of genuine and imposter scores in proposed MSS

(ii) For face identification experiments, fa set is used as gallery set, while dup I and dup II are used as probe sets. In case of weighted DSAF weights pertaining to each region are assigned using the individual recognition accuracies of region 1, region 2, region 3 and region 4 of difference half face images as recorded in Table 5.2, for the above mentioned experimental setup. The performance of proposed combined, DSAF, and weighted DSAF methods is given by ROC and CMC curves for dup I and dup II sets as shown in Figures 5.9 and 5.10, respectively.

5.4.2 Experiments on MORPH Database

(i) For face verification experiments, a subset consisting of 40000 face images of 10000 subjects from MORPH database is selected. Starting from youngest face image of each subject, we select 2 images per subject in training resulting into 10000 genuine and 99990000 imposter matching scores in training, each for PCA, LBP, and weighted DSAF

algorithms. 10000 older and 10000 oldest images are used as probe sets for face verification against small and large temporal variations, respectively.

(ii) For face identification experiments, the youngest face images set is used as gallery set, while older and oldest face images sets are used as probe sets. For weighted DSAF, weights pertaining to each region are assigned using the individual accuracies of four regions of difference half face images as given in Table 5.2 for the above mentioned experimental setup. The performance of proposed combined, DSAF, and, weighted DSAF methods is given by ROC and CMC curves for older and oldest image sets, as shown in Figures 5.11 and 5.12, respectively.

5.4.3 Experiments on FG-NET Database

(i) For face verification experiments, a subset consisting of 328 face images of 82 subjects from FG-NET database is selected. Starting from the youngest face image of each subject, 2 images per subject are selected, resulting into 82 genuine and 6642 imposter matching scores in training, each for PCA, LBP, and weighted DSAF algorithms. 82 older and 82 oldest images are used as probe sets for face verification against small and large temporal variations, respectively.

(ii) For face identification experiments, the youngest face images set is used as gallery set, while older and oldest face images sets are used as probe sets. In case of weighted DSAF, weights pertaining to each region are assigned using the individual accuracies of four regions of difference half face images as recorded in Table 5.2 for the same experimental set up as described above. The performance of proposed combined, DSAF, and weighted DSAF methods is given by ROC and CMC curves for older and oldest image sets, as shown in Figures 5.13 and 5.14, respectively.

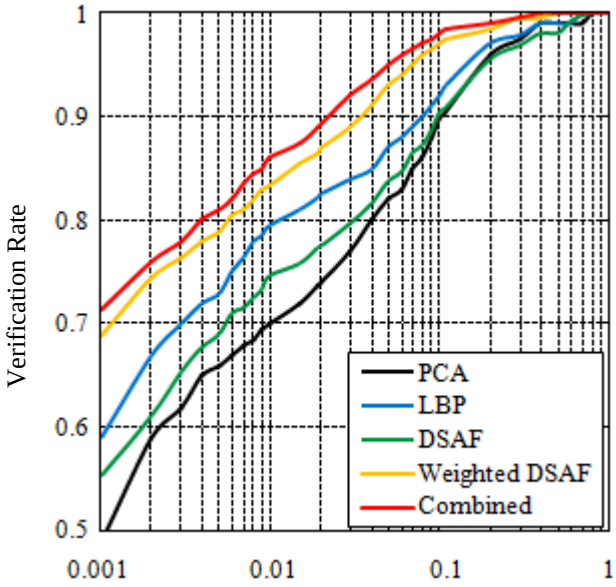
The individual recognition accuracies of asymmetric facial regions show that region 4 gives the highest recognition accuracy compared to the other regions for selected sub sets of face images from FERET, MORPH, and FG-NET databases. This fact can be attributed to most asymmetric facial information present in region 4, which is in line with facial asymmetry evaluation results suggesting an increase in facial asymmetry for lower parts compared to upper in middle parts of face images.

Table 5.2: Individual recognition accuracies of region 1, region 2, region 3 and region 4 of difference half face images for FERET, MORPH, and FG-NET databases

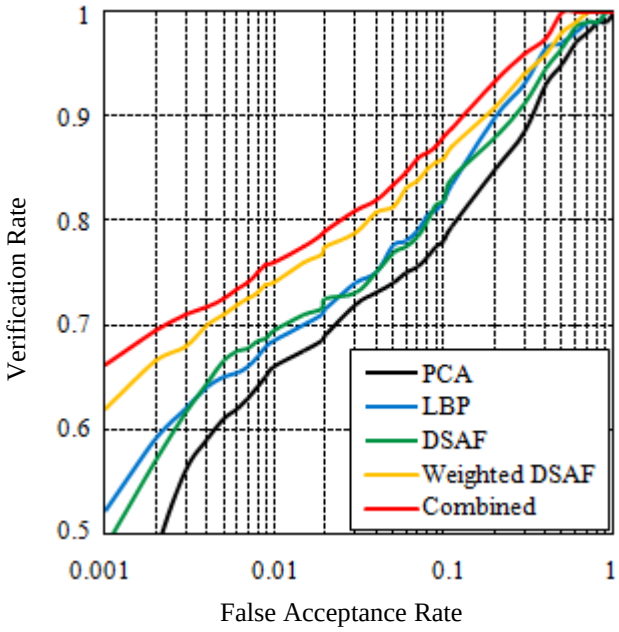
Database	Probe set	Verification accuracy @ 0.001 FAR (%)				Rank-1 identification accuracy (%)			
		Region 1	Region 2	Region 3	Region 4	Region 1	Region 2	Region 3	Region 4
FERET	dup I	31.99	34.62	36.98	41.55	29.77	32.68	33.10	38.64
	dup II	29.91	31.02	32.96	37.11	25.20	27.00	31.44	36.01
MORPH	Older images	33.20	35.00	38.33	42.00	31.50	32.88	36.20	39.50
	Oldest images	30.58	32.60	34.90	43.60	30.00	31.60	33.00	37.40
FG-NET	Older images	31.70	34.14	36.58	42.68	30.48	35.36	37.80	40.24
	Oldest images	29.26	31.70	32.92	41.46	28.04	32.92	36.58	39.02

5.4.4 Comparison with Existing Algorithms

The performance of proposed combined, DSAF, and weighted DSAF approach is compared with individual algorithms including PCA and LBP, as shown in Figures 5.10, 5.12, and 5.14. The performance of proposed methods is also compared with the following two state-of-the-art algorithms, developed for recognition of age-separated face images. (i) Bacteria foraging fusion (BFF) [13]: In BFF, LBP descriptor is used to calculate matching scores for different facial parts (mouth, binocular, and periocular). Weights pertaining to different facial parts are then optimized to enhance the recognition performance of age-separated face images. (ii) Multi-view discriminative learning (MDL) [62]: In MDL, three local facial features, LBP, scale invariant feature transform (SIFT), and gradient orientation pyramid (GOP), are extracted from overlapping patches of face images. An optimization problem is then solved to maximize inter-class separation and minimize intra-class separation to recognize age-separated face images. The Rank-1 recognition performance of proposed methods is compared with existing methods in the form of CMC curves as shown in Figures 5.10, 5.12, and 5.14, both for small and large temporal variations on FERET, MORPH, and FG-NET databases, respectively.

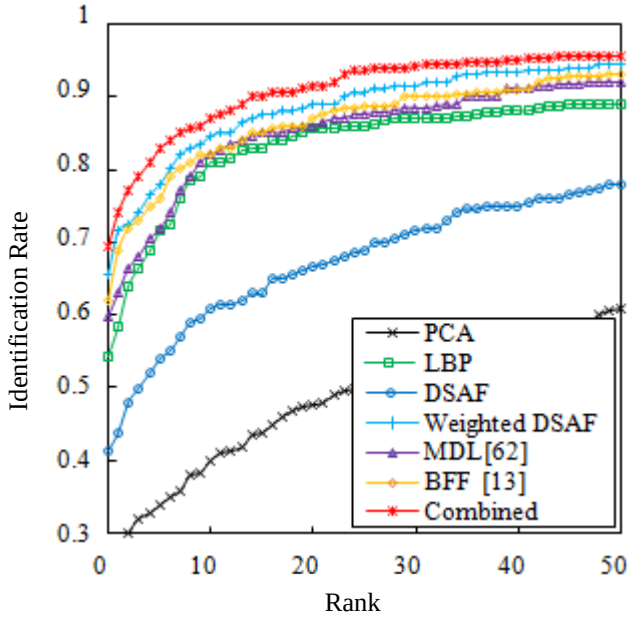


(a)

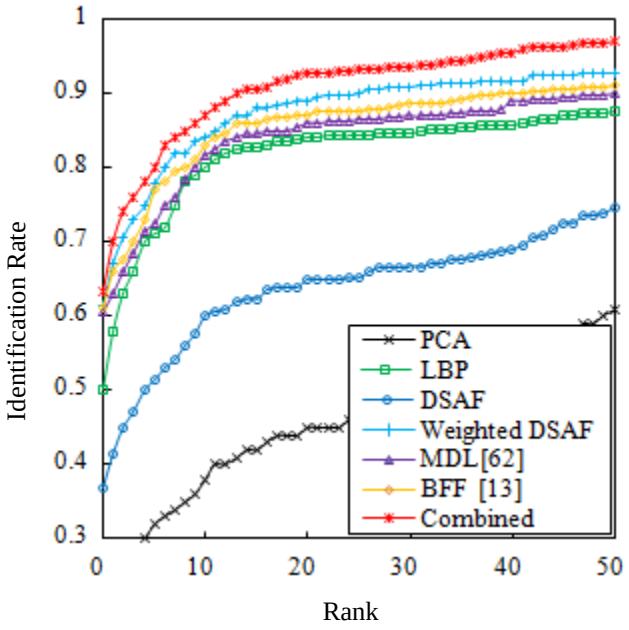


(b)

Figure 5.9: ROC curves showing the performance of proposed combined, DSAF, and weighted DSAF methods, against existing methods for (a) dup I and (b) dup II sets on FERET database

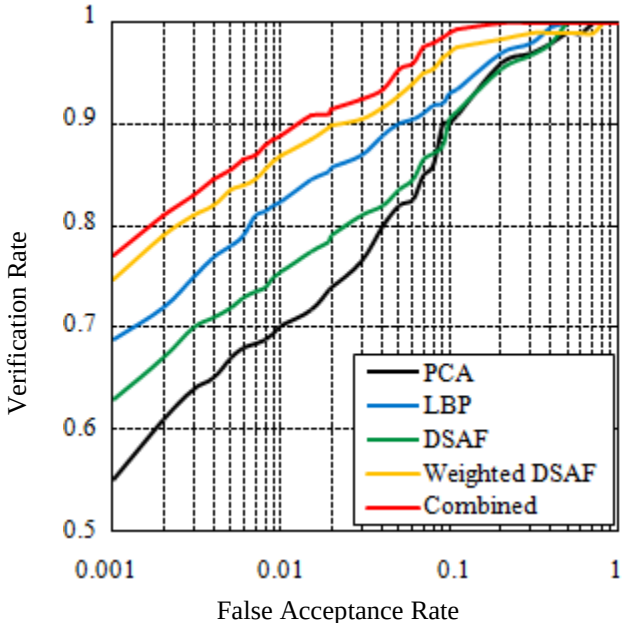


(a)

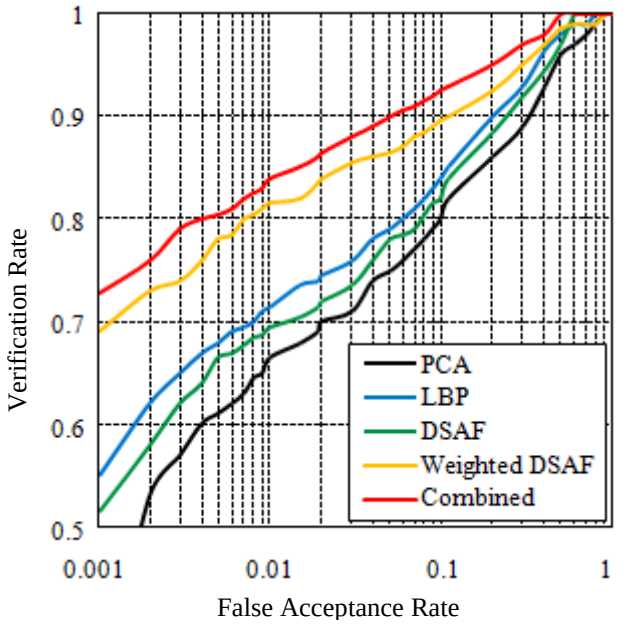


(b)

Figure 5.10: CMC curves showing the performance of proposed combined, DSAF, and weighted DSAF methods against existing methods for (a) dup I and (b) dup II sets on FERET database

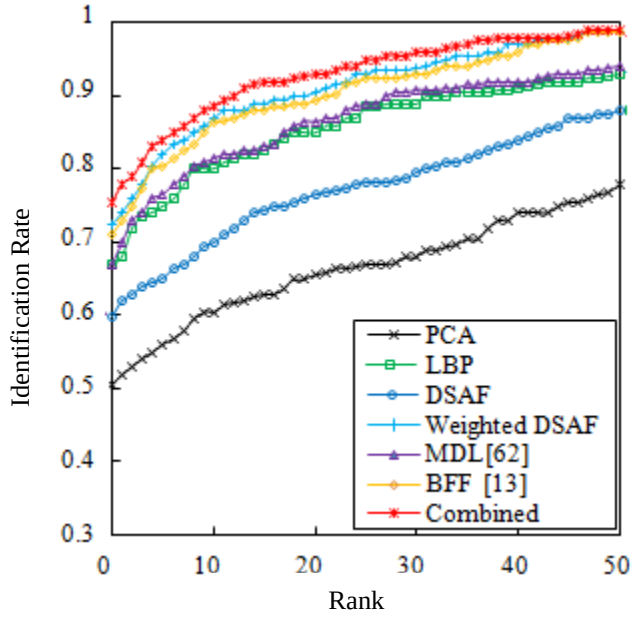


(a)

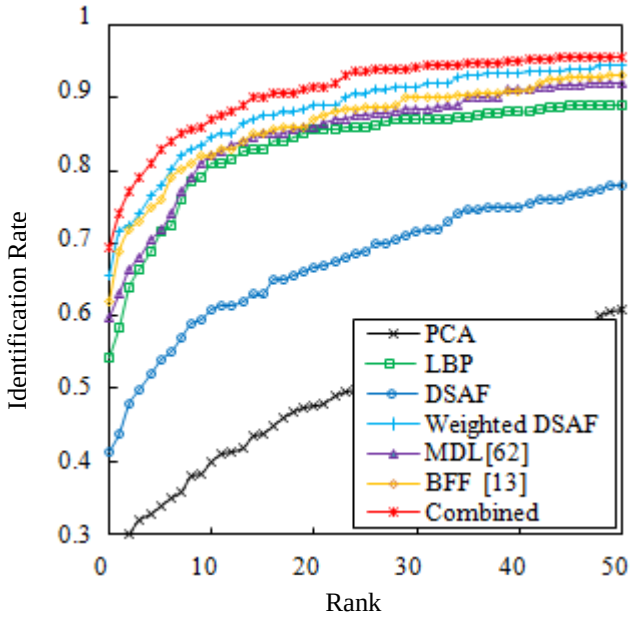


(b)

Figure 5.11: ROC curves showing the performance of proposed combined, DSAF, and weighted DSAF methods, against existing methods for (a) older and (b) oldest image sets on MORPH database

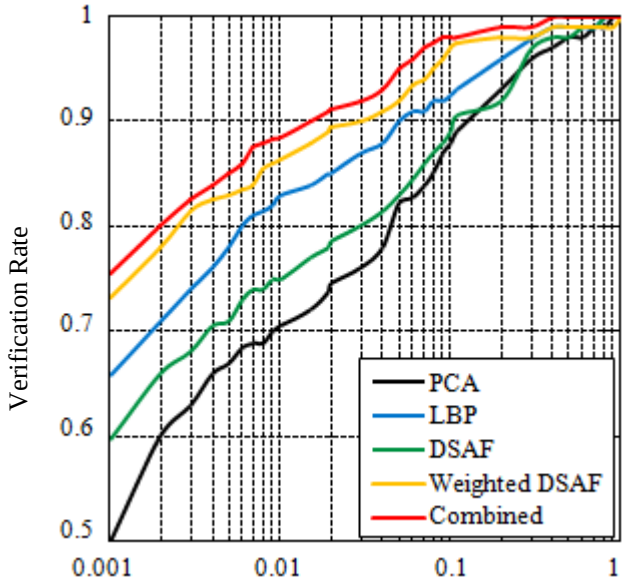


(a)

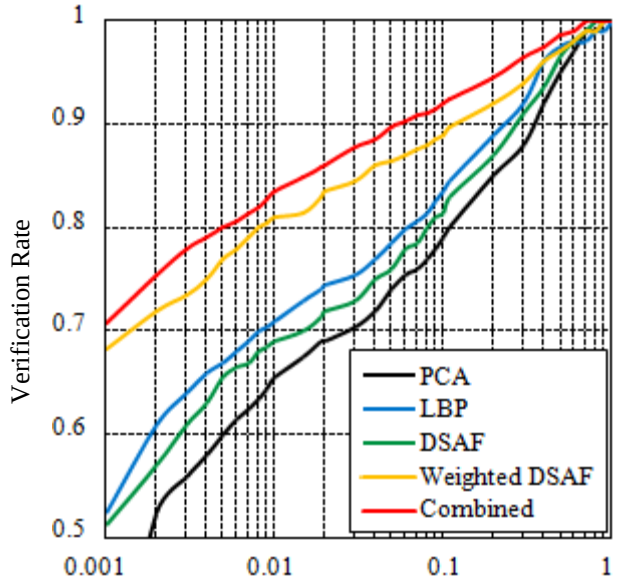


(b)

Figure 5.12: CMC curves showing the performance of proposed combined, DSAF, and weighted DSAF methods, against existing methods for (a) older and (b) oldest face images on MORPH database

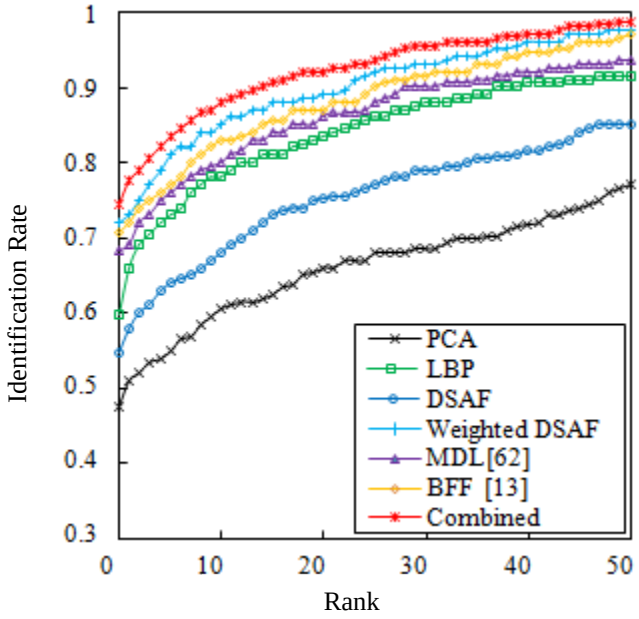


(a)

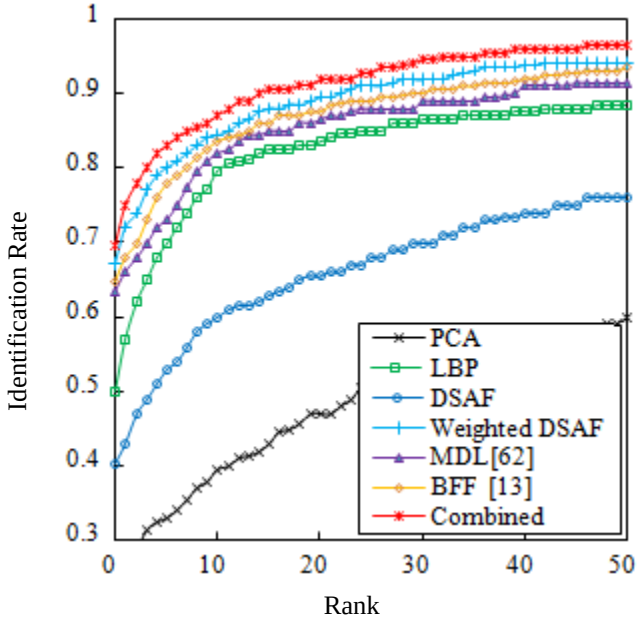


(b)

Figure 5.13: ROC curves showing the performance of proposed combined, DSAF, and weighted DSAF methods, against existing methods for (a) older and (b) oldest image sets on FG-NET database



(a)



(b)

Figure 5.14: CMC curves showing the performance of proposed combined, DSAF, and weighted DSAF methods, against existing methods for (a) older and (b) oldest image sets on FG-NET database

Table 5.3 summarizes face verification accuracies at 0.001 FAR and the Rank-1 identification accuracies for proposed and existing methods for FERET, MORPH, and FG-NET databases on two protocols: the image sets with (i) small temporal variations and (ii) large temporal variations are used as probe sets.

Table 5.3: Recognition accuracies of proposed combined, DSAF, and weighted DSAF methods vs. existing methods

Database	Probe set	Verification accuracy at 0.001 FAR (%)					Rank-1 identification accuracy (%)						
		PCA	DSAF	LBP	Weighted DSAF	Proposed combined	PCA	LBP	DSAF	Weighted DSAF	BFF [13]	MDL [62]	Proposed combined
FERET	Dup I	49.03	55.40	59.00	69.11	**71.32**	44.04	62.04	56.92	68.97	68.42	67.03	**70.08**
	Dup II	29.91	48.71	52.13	61.96	**66.23**	21.79	50.00	36.75	61.53	61.11	60.68	**63.24**
MORPH	Older images	55.00	63.00	68.80	74.70	**77.20**	50.51	66.89	60.00	72.35	71.00	67.00	**75.40**
	Oldest images	32.40	51.50	55.00	69.00	**72.80**	26.00	54.19	41.50	65.50	62.10	59.90	**69.40**
FG-NET	Older images	50.00	59.75	65.85	73.17	**75.60**	47.56	59.75	54.87	71.95	70.73	68.29	**74.39**
	Oldest images	31.70	51.21	52.43	68.29	**70.73**	25.60	50.00	40.24	67.07	64.63	63.41	**69.51**

5.4.5 Computational Complexity Analysis

The computational complexity of proposed combined method is analyzed in terms of feature extraction, matching scores calculation, and classification stage. We assume that number of pixels and overlapping patches in face image is *N* and *xy*, while for difference half face image is *N* and *XY*, respectively. The computational complexity of feature extraction algorithms (PCA, LBP, and weighted DSAF) is given in Table 5.4.

The computational complexity of matching scores calculation stage has an order of $O(s^2)$, where s represents the number of matching scores calculated for PCA, LBP, and weighted DSAF feature vectors. Finally, the computational complexity of SVM classifier with RBF kernel is *O(KNsv)* where *K* and *Nsv* represent number of testing samples and support vectors, respectively.

Table 5.4: Computational complexity of feature extraction algorithms for proposed combined method

Algorithm	PCA	LBP		Weighted DSAF		
				APs	Hu moment invariants	
Complexity	For face	For patch	For face	For difference half face	For patch	For difference half face image with extracted asymmetries
	$O(VET)^{\dagger}$	$O(N_d)$	$O(xyN_d)$	$(GN_d+4LN_d)^{\ddagger}$	$O(N^2_d)$	$O(XYN^2_d)$

† V, E, and T represent the number of projection vectors, the testing and training samples, respectively.

‡ G and L represent number of gray levels in the difference half face image and number of transformation levels of APs, respectively.

The complexity analysis shows that proposed algorithm is linear with respect to PCA, LBP, and APs along with SVM classifier. The most demanding stages are computation of Hu moment invariants and calculation of matching scores, with upper bound of $O(N^2_d)$ and $O(s^2)$, respectively.

5.5 Results Related Discussion

Face verification performance: The verification results show that it is more difficult to recognize face images with large temporal variations compared to those with small temporal variations. The proposed combined and weighted DSAF approaches achieve superior performance compared to PCA, LBP, and DSAF. At 0.001 FAR, the verification accuracies of 71.32% and 66.23% on FERET, 77.20% and 72.80% on MORPH, while 75.60% and 70.73% on FG-NET database are achieved for small and large temporal variations, respectively. Thus proposed combined approach boosts the recognition performance of age-separated face images, considerably.

Face identification performance: The proposed combined and weighted DSAF approaches achieve better Rank-1 recognition accuracies compared to PCA, LBP, and DSAF. We achieve an accuracy of 70.08% and 63.24% on FERET, 75.40% and 69.40% on MORPH while 74.39% and 69.51% on FG-NET database for small and large temporal variations, respectively. Thus a considerable increase in identification performance is achieved by combined approach. Performance of proposed methods is also compared with two recent algorithms, BFF and MDL, developed for recognition of age-separated

face images. In all identification experiments, it is observed that proposed combined approach outperforms the existing methods. MDL gives 67.03% and 60.68% on FERET, 67.00% and 59.90% on MORPH while 68.29% and 63.41% on FG-NET database for small and large temporal variations, respectively. BFF gives 68.42 % and 61.11% on FERET, 71.00 % and 62.10% on MORPH while 70.73% and 64.63% on FG-NET database for small and large temporal variations, respectively. Compared to the proposed combined approach, BFF and MDL rely only on the local facial features which, in our view, can be manipulated by the extrinsic facial variations like the facial make-up. In the current study, the holistic and local facial features have been combined with asymmetric facial features which are difficult to be manipulated by such facial variations.

5.6 Summary

In this chapter a facial asymmetry based approach has been presented to recognize age-separated face images. Each face image was represented by extracting existing PCA and LBP based features along with newly proposed DSAF features. Matching scores of respective feature vectors were then combined in proposed matching-score space (MSS) to discriminate between genuine and imposter classes, using Support Vector Machine (SVM) as a classifier. In conclusion, firstly it is observed that the facial asymmetry is an intrinsic facial feature, which can be used to recognize age-separated face images. Secondly, it is deduced that the proposed matching-score space based approach yields better recognition performance compared to that of individual matching scores of PCA, LBP, and DSAF features. Thirdly, the proposed approach is more adaptable to recognize age-separated face images, compared to some existing state-of-the-art methods. Finally, the recognition results suggest that it is more difficult to recognize face images with large temporal variations compared to those with small temporal variations.

Next we plan to use some of these insights in developing a framework capable of estimating the age group of a given query face image before recognizing it across temporal variations. We hope that integrating the knowledge learned from age group estimation into a face recognition algorithm will enhance the recognition performance considerably.

Chapter 6

AGE GROUP ESTIMATION BASED RECOGNITION

In this chapter we introduce the problem of the facial asymmetry based age group estimation and its role in recognizing age-separated face images. Age group estimation plays an important role in recognizing age-separated face images. Age group estimation also finds numerous applications in different computer vision tasks. Here a framework is proposed to integrate the knowledge learned from age group estimation into a face recognition algorithm to enhance the recognition performance.

6.1 Introduction

Accurate age group estimation finds numerous applications such as statistical analysis (for example, class-wise age distribution), limited access to purchase of certain commodities (for example, alcohol, and tobacco) and human-computer interaction (for example, limited internet access to certain age groups). Being an active research area, both the human perceptions based and machine based age group estimation algorithms have been reported in literature. Despite significant advances and ample work in related research area [28], [41], [45], [50], [51], [52], [86], yet existing methods have certain limitations including: (i) use of such discriminative features which are vulnerable to be manipulated by certain extrinsic variations like facial makeup [16], (ii) there is lack of such related studies which integrate the knowledge learned from age group estimation into face recognition algorithms. In this chapter, asymmetric facial features are used for age group estimation task. In this research, we have focused on estimating the age groups rather than the exact age. This can potentially help in learning important characteristics such as features for every age group, which can be utilized during face recognition. We also devise a method to integrate the knowledge learned from age group estimation into face recognition algorithm to enhance the recognition performance.

6.2 Age Group Estimation

6.2.1 Pre-Processing

For age group estimation task, we use the benchmark MORPH and FERET facial aging databases. In order to mitigate the effects of different appearance variations, all the face images are pre-processed as described in following four steps.

(i) All the face images are rotated and aligned, based on eye coordinates such that these are vertically upright.

(ii) All the face images are then cropped into 128×128 pixel size with an equal inter-pupillary distance (IPD).

(iii) Colour face images are converted to grey scale images in order to alleviate the unwanted colour effects. For this purpose we use a luminance model adopted by NTSC and JPEG: $Y = 0.299R + 0.587G + 0.114B$, where R, G and B represent Red, Green and Blue colour channels and Y is resulting grey scale image.

(iv) To remove unwanted illumination variations from resulting grey scale images, histogram equalization is used.

After the pre-processing stage, facial landmarks are detected using a state-of-the-art face recognition and analysis tool, Face++ [26]. Sixty-four facial landmarks are detected for each face image as shown in Fig. 6.1.

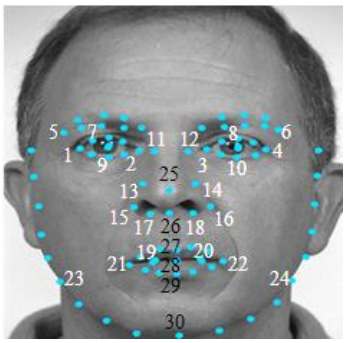


Figure 6.1: Automatic facial land mark detection using Face++ [25]

We select 30 landmarks (shown by numbers in Fig. 6.1) to extract twelve horizontal (HFA) and eight vertical (VFA) facial asymmetric distances (termed as dimensions in

rest of text) of selected pairs of bilateral facial land marks. Horizontal distances are measured reference to facial mid line y_1y_2 , and line x_1x_2 joining the eye coordinates (see Figure 5(a), Chapter 5).The resulting horizontal and vertical asymmetric distances are given in Table 6.1 and Table 6.2, respectively.

Table 6.1: Horizontal dimensions

Pairs of bilateral facial landmarks	1,4	2,3	5,6	7,8	9,10	11,12	13,14	15,16	17,18	19,20	21,22	23, 24
Horizontal dimensions	HA1	HA2	HA3	HA4	HA5	HA6	HA7	HA8	HA9	HA10	HA11	HA12

Table 6.2: Vertical dimensions

Pairs of bilateral facial land marks	9,10	11,12	13,14	15,16	17,18	19,20	21,22	23,24
Vertical dimensions	VA1	VA2	VA3	VA4	VA5	VA6	VA7	VA8

We use the selected horizontal and vertical dimensions to assess the correlation between horizontal and vertical dimensions; and finally develop a facial asymmetry based hierarchical approach to perform age group estimation, as described in the following sub-sections.

In order to analyze the mutual independence of horizontal and vertical dimensions, we perform Pearson correlation test between twelve horizontal and eight vertical dimensions. We observe that 83.33% of the horizontal and vertical dimensions represent weak correlation (less than 0.5) and hence are mutually independent. Motivated by this fact we will use both the horizontal and vertical dimensions as independent measures of facial asymmetry in subsequent analysis for age group estimation, as described in the following section.

6.2.2 A Hierarchical Approach for Age Group Estimation

In this section, a hierarchical approach is presented for age group estimation task. For this purpose, face images are normalized against different types of appearance variations as explained in previous section. In order to perform facial asymmetry based age group estimation, it is desirable to represent face images in terms of horizontal and vertical

dimensions and select a subset of these dimensions such that selected subset can discriminate between images belonging to two different age groups. The F-test [87] is used for dimension selection, while the Canonical Correlation Analysis (CCA) [88], [89] is used to find a linear combination of selected subset of dimensions and hence calculate discriminating score to classify the face images belonging to two different age groups. Finally, the discriminating scores are applied as an input to a Support Vector Machine (SVM) based binary tree classifier to achieve age group estimation. An overview of the proposed approach is shown in Figure 6.2.

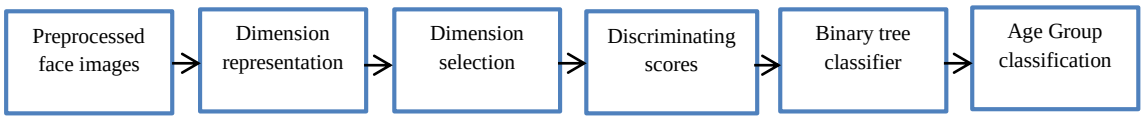


Figure 6.2: Overview of the proposed age group estimation approach

Dimensions selection: The qualitative analysis of the projection of original 12 horizontal and 8 vertical dimensions on to first four principal components shows that it is not possible to discriminate face images beloning to two given age groups based on these 20 dimensions altogether. This outcome reveals the fact that original 20-dimensional space is unable to discriminate between two age groups, based on their simple linear combinations. Hence a statistical procedure is proposed to determine if a subset of original dimensions can discriminate between two age groups for a given set of face images, as illustrated below.

Given a training set G with s samples such that

$$G = \{(x_i, y_i) : \ x_i \in R^v, y_i \in N, i \in [1, s]\} \tag{6.1}$$

where x_i represent v-dimensional feature vectors, representing original dimensions from i^{th} training face image and y_i represent corresponding labels. Our purpose is to select a subset H with dimension v' of original dimensions such that

$$H = \left\{x_i' : x_i' \in R^{v'}, \ \ x_i' \subset x_i, \ \ i \in [1, s]\right\} \tag{6.2}$$

where $v' < v$, which can retain discriminative dimensions to classify face images belonging to two given age groups. For age group classification task, the labels $y_i \in \{0, 1\}$, represent the respective age groups i.e. {age group 1, age group 2}. In order to accomplish this, a given set of face images is divided into two groups, with a set of 20 original dimensions (12 horizontal and 8 vertical) extracted for each face image. The univariate F-test [88] is performed for each of the 20 original dimensions to measure the significant difference between the means of dimensions of two age groups. The F-test for a particular dimension can be defined as shown in equation (6.3).

$$F = \frac{Total\ variance}{Average\ within\ group\ variance} \tag{6.3}$$

The larger value of *F* for a given dimension shows that it is more discriminating between two age groups, compared to the other dimensions. The procedure is repeated for all dimensions and only those dimensions are selected whose *F* values greater than a pre-determined threshold, determined empirically. All other dimensions are discarded and further analysis is performed using the selected dimensions only. We perform the F-test on MORPH and FERET datasets as described below.

10000 youngest face images of 10000 subjects are selected from MORPH database (one image per subject) and 1196 face images of 1196 subjects (one image per subject) are selected from fa set of FERET database. The F-test is performed on seleted image sets by dividing images into two age groups, age group 1 (16-30 for MORPH and 10-30 for FERET datasets) and age group 2 (31-60+ in case of both MORPH and of FERET datasets), based on ground truth information. Table 6.3 illustrates the F-test results for two age groups for MORPH and FERET datasets, with top five selected dimensions (HA1, HA2, HA3, VA1, and VA2) along with corresponding *p* values, suggesting the significance of selected dimensions. By comparision, it is observed that these five dimensions (termed as discriminating dimensions in the rest of text), are most discriminating among other dimensions to discriminate between two selected age groups.

In order to assess the robustness of proposed approach, the above mentioned procedure is repeated for older and oldest face image sets from MORPH and dup I and dup II sets from FERET datasets. The rate of recurrence of each measurement is recorded for each of 3 image sets, as shown in Figures 6.3 (a) and 6.3 (b), respectively. The results show the most frequent occurance of HA1, HA2, HA3, VA1, and VA2, suggesting the robustness of proposed approach in selecting the most discriminating dimensions for age group classification.

Table 6.3: F-test results for top 5 discriminating dimensions for MORPH and FERET datasets

	MORPH					FERET				
Selected dimensions	HA1	HA2	VA2	HA3	VA1	HA1	HA2	VA2	HA3	VA1
F values	7.0	4.1	2.5	1.8	1.2	6.6	3.2	2.0	1.1	1.0
P values	0.001	0.000	0.003	0.002	0.001	0.000	0.000	0.004	0.001	0.002

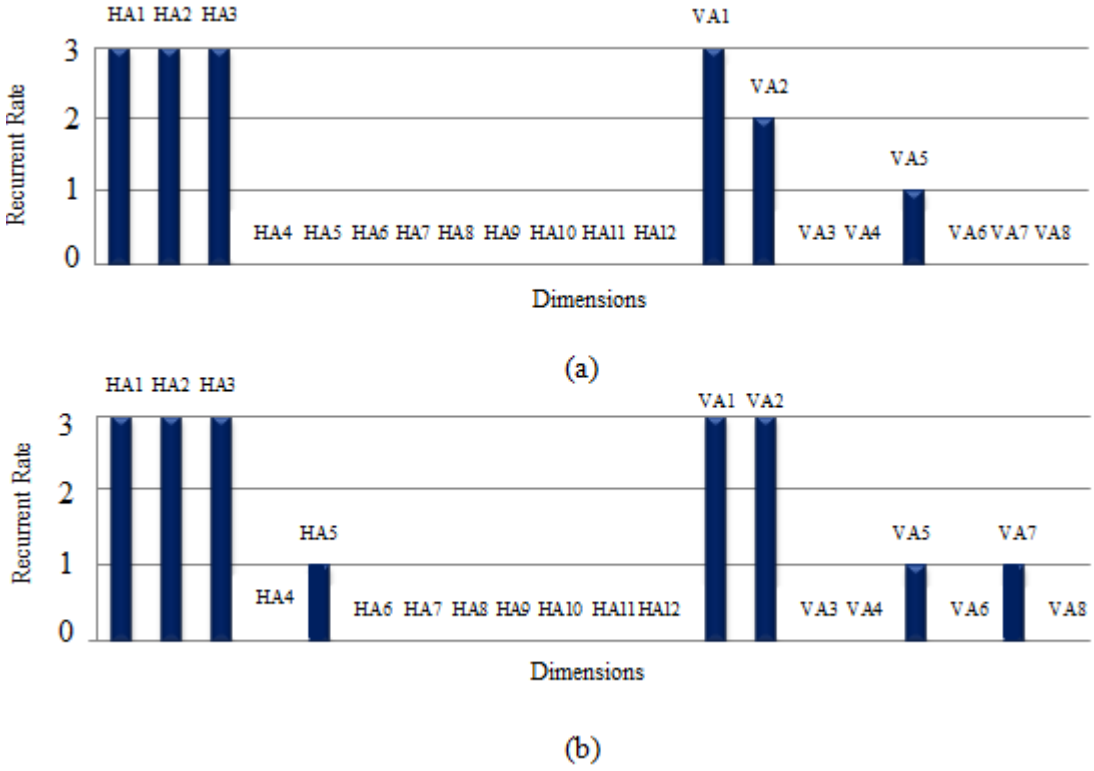


Figure 6.3: Rate of recurrence of selected dimensions extracted from three different image sets from (a) MORPH and (b) FERET database

The same procedure is repeated by splitting the selected image sets from MORPH and FERET datasets into four age groups, age group 11, age group 12, age group 21 and age group 22 (16-20, 21-30, 31-45 and 46-60+ for MORPH, while 10-20, 21-30, 31-45 and 46-60+ for FERET dataset, respectively). It is observed that dimensions HA1, HA4, HA7, VA1 and VA3 are most discriminating between age group 11 and age group 12, while the dimensions HA2, HA7, VA1, VA3, and VA6 are most discriminating between age group 21 and age group 22, both for MORPH and FERET datasets.

Discriminating scores: After the selection of a subset of most discriminating dimensions through F-test, we aim at learning (i) a subspace, based on linear combination of discriminating dimensions to calculate the discriminating scores to classify face images into two age groups; and (ii) the relative contribution of each discriminating dimension to classify two age groups. For this purpose, the well-known Canonical Correlation Analysis (CCA) is used as described below.

Canonical correlation analysis (CCA): For a given set of objects with two different representations, Canonical correlation analysis (CCA) aims at computing a projection such that, the correlation between two representations is maximum in a sub-space with reduced dimensionality. Given a set of observations, $(U, V) = \{(u^{(j)}, v^{(j)}\}_{j=1}^{n}$ CCA seeks to find pairs of projection directions $\{(x_i, y_i)\}_{i=1}^{d}$, such that the correlation between projected inputs and outputs is maximized, as given in expression (6.4).

$$\{(x_i, y_i)\}_{i=1}^{d} \leftarrow CCA(U, V) \tag{6.4}$$

In case of age group classification, it is observed that a projection direction with the largest eigen value is significantly higher than the second largest eigen value, verifying the possibility to represent and hence discriminate the data in a single dimension. We also calculate the relative contribution for each discriminating dimension in classifying the two age groups, by calculating the magnitude and polarity of relevant canonical coefficients. The linear combination of such dimensions then results into a linear function representing the face images from two age groups. For a set of 200 face images from youngest face images set of MORPH dataset, sample canonical coefficients with relative

magnitude and polarity are shown in Table 6.4. It is observed that larger the values of HA1 and VA1, more likely is that the subject belongs to age group 1. Similarly larger the values of HA2, HA3 and VA2, the subject will belong to age group 2. A sample scatter plot showing the discriminating scores for two age groups (16-30 and 31-60+) for 200 youngest face images set from MORPH dataset is shown in Figure 6.4.

Table 6.4: Relative contributions of discriminating dimensions determined using CCA

Selected dimensions	HA1	HA2	HA3	VA1	VA2
Canonical coefficients (magnitude and ploarity)	-12.05	3.6	7.9	-6.0	2.5

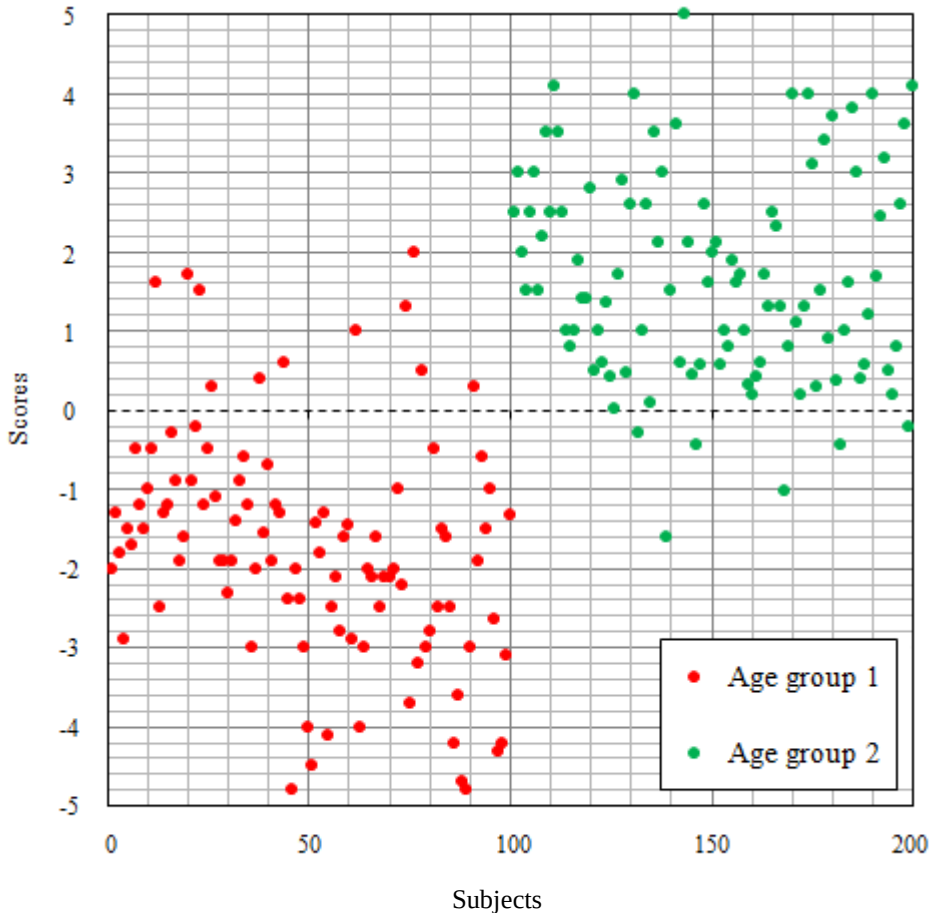


Figure 6.4: A sample scatter plot showing discriminating scores for two age groups of 200 subjects from MORPH dataset

From scatter plot, it is obvious that the discriminating scores determined through CCA cannot completely distinguish between two age groups due to overlapping. In order to improve the classification, the discriminating scores are applied to a SVM based binary tree classifier, as explained in the following subsection.

Binary tree classification for age group estimation: The proposed age group estimation problem can be defined as a binary classification task. To handle such classification problem, a hierarchical method is proposed, which uses DSAF feature descriptor as an input to SVM based binary tree classifiers, as shown in Figure 6.5. Three binary SVM classifiers (C, C_1, and C_2) are used to build a two-level binary decision tree [90].The labeled discriminating scores obtained through CCA algorithm are applied as an input to learn a binary tree classifier such that a given face image can be classified into one of the age groups (age group 1 or age group 2) at level 1 and into one of four age groups (age group 11, age group 12, age group 21 or age group 22) at level-2 of binary tree classifier.

Once trained, a given test face image can be classified into one of the binary age groups (age group 1 or age group 2) at level-1 or into one of the four age groups (age group 11, age group 12, age group 21 and age group 22) at level-2 of binary decision tree.

Support Vector Machine has been successfully used for age group classification problems as reported in [28]. In the proposed binary tree classifier, a publically available LIBSVM source [85] is used for implementation. We use the labels $0,1$ for binary classification of two age groups while training the SVM based binary tree classifiers, with RBF kernel. The values of parameters c and γ are selected by using a 5-fold cross validation protocol on training set. The thresholds t, t_1, and t_2 used to partition the age range, are empirically determined.

Hierarchical methods have been studied for age group estimation problems in [47], [91] but our approach differs from these methods such that instead of dividing the entire age-range directly into certain number of age groups (for example four age groups for the age

range of 0-70 years in [91]), our approach uses a coarse-to-fine strategy for age group classification, similar to that proposed in [50].

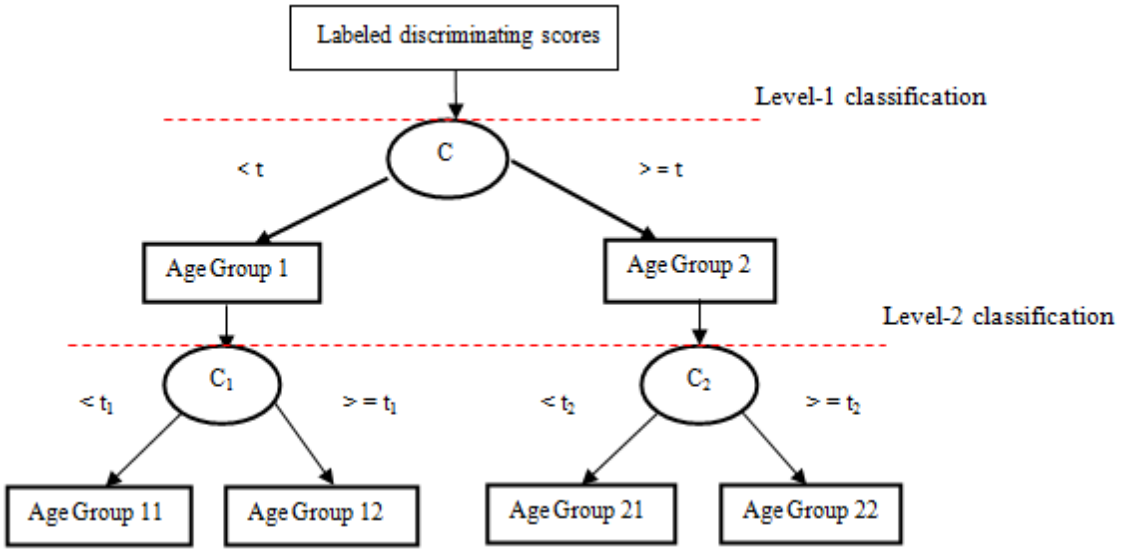


Figure 6.5: Learning a 2-level hierarchical age group estimator based on SVM binary tree classifiers

6.2.3 Experimental Results of Age Group Estimation Algorithm

To evaluate the performance of proposed age group estimation approach, 10000 youngest face images from MORPH database is partitioned into training and a test set consisting of 5000 face images each with age range of 16-60+ years. Similarly 1196 face images from fa set of FERET database is partitioned into training and a test set, comprising of 598 face images each, with age range of 10-60+ years. A five-fold cross validation method is used to evaluate the proposed age group estimation algorithm.

The age group estimation performance is reported in a trained-predicted confusion matrix (see Table 6.5) with age of database subjects belongs to one of four age groups 16-20, 21-

30, 31-45 and 46-60+ for MORPH, while 10-20, 21-30, 31-45, and 46-60+ in case of FERET dataset.

Table 6.5: Confusion matrix for age group classification accuracy for (a) MORPH and (b) FERET dataset

Training age groups	Predicted age groups			
	16-20	21-30	31-45	46-60+
16-20	**748**	58	10	6
21-30	20	**1349**	40	11
31-45	5	30	**2048**	50
46-60+	11	9	21	**584**

(a)

Training age groups	Predicted age groups			
	10-20	21-30	31-45	46-60+
10-20	**211**	11	7	5
21-30	2	**122**	3	3
31-45	0	1	**80**	3
46-60+	1	3	7	**139**

(b)

An important observation from the age group estimation performance is the least confusion of age group 31-45 and 46-60+ with corresponding younger age groups, as opposed to the trend reported in [28] where subjects from LFW+ database in a similar age range tend to be confused with younger age groups, due to possible use of facial make up. In proposed study, we attribute this least confusion to asymmetric facial features which are difficult to manipulate by extrinsic facial variations like facial makeup.

6.3 Comparison with Existing Methods

For the sake of comparison, we evaluate the age group estimation performance of the following state-of-the-art methods on same image sets used in proposed method.

(i) A state-of-the-art method based on variants of Local Binary Patterns (LBP) presented in [51].
(ii) An independently trained deep face representation algorithm, Face++ [25].
(iii) Random Forest Classifier (RFC), an ensemble of trees based method [92].

Sensitivity as a measure of overall accuracy is computed for proposed and existing methods, representing the true positive performance [93]. The variants of LBP [51] give an overall accuracy of 79.41% and 77.99% on MORPH and FERET datasets, respectively.

Face++ [25] gives an overall age group estimation accuracy of 87.26% and 86.23% on MORPH and FERET datasets, respectively.

The RFC gives an overall age group accuracy of 83.15% and 85.23% on MORPH and FERET datasets, respectively compared to proposed method which gives an overall accuracy of 93.86% and 92.98% on MORPH and FERET datasets, respectively showing the superior performance of the proposed method to estimate the age groups of given face images.

Table 6.6 summarizes the age group estimation accuracies both for MORPH and FERET datasets for proposed and existing methods.

Table 6.6: Age group estimation performance comparison with state-of-the-art methods for (a) MORPH (b) FERET dataset

Method	Age groups				Sensitivity or overall accuracy (%)
	16-20	21-30	31-45	46-60+	
Variants of LBP [52]	81.99	80.00	76.98	78.67	79.41
Face++ [26]	86.49	86.97	90.62	84.96	87.26
RFC [93]	87.20	84.90	81.50	79.00	83.15
Proposed	90.99	95.00	96.01	93.44	93.86

(a)

Method	Age groups				Sensitivity or overall accuracy (%)
	10-20	21-30	31-45	46-60+	
Variants of LBP [52]	80.76	77.69	76.19	77.33	77.99
Face++ [26]	85.47	86.15	89.28	84.00	86.23
RFC [93]	86.50	86.11	84.33	84.00	85.23
Proposed	90.17	93.84	95.23	92.67	92.98

(b)

6.4 Effect of Asymmetric Facial Regions on Age Group Estimation

We notice variation of facial asymmetry for different facial landmarks while evaluating facial asymmetry in chapter 5. Motivated by this fact, we are interested in understanding the effect of various asymmetric facial regions on age group estimation. For this purpose, each difference half face with extracted asymmetries is divided into four regions, region 1 region 2, region 3, and region 4. We use each region separately to estimate the age groups using similar experimental set up used in the age group estimation task described earlier.

Table 6.7 represents the results for this experiment, both for MORPH and FERET datasets.

Table 6.7: Age group estimation accuracy for different facial regions on (a) MORPH (b) FERET dataset

	Asymmetric facial regions accuracy (%)			
Age group	Region 1	Region 2	Region 3	Region 4
16-20	37.95	43.66	44.02	46.56
21-30	35.03	39.50	40.03	40.96
31-45	31.99	41.47	38.02	42.08
46-60+	27.61	37.53	38.02	40.48

(a)

	Asymmetric facial regions accuracy (%)			
Age group	Region 1	Region 2	Region 3	Region 4
10-20	38.03	43.07	45.23	47.33
21-30	33.33	38.46	39.28	40.00
31-45	30.34	36.15	37.90	39.33
46-60+	26.06	36.15	38.09	41.33

(b)

By comparison, it is observed that estimation accuracy increases from upper to lower regions of face, which can be attributed to more pronounced facial asymmetry for these regions and hence more prominent role of corresponding dimensions in age group estimation. For example in case of age group 16-20 for MORPH dataset, the region 4 (comprising of lower facial landmarks) gives an age group estimation accuracy of 46.56% compared to region 1 with an age group estimation accuracy of 37.95% (comprising of upper facial landmarks). Thus we conclude that more asymmetric a facial region is, more it contributes in estimating age group of given face image.

6.5 Recognition of Age-Separated Face Images

After assessment of facial asymmetry across age variations and its potential role in age group estimation, the next step is to design an algorithm which can recognize age-separated face images based on knowledge learned from age group estimation of query face image. Before we implement such an algorithm, we are interested in understanding the role of various facial asymmetric regions in recognizing age-separated face images. For this purpose densely asymmetric facial features (DSAF) are used as introduced in chapter 4.

6.5.1 Effect of Asymmetric Facial Regions on Recognition Accuracy of Age-Separated Face Images

In order to understand the effect of various asymmetric facial regions in recognizing age-separated face images, face identification experiments are performed by using MORPH and FERET training and probe sets. 10000 youngest face images are used as training set while older and oldest face image sets with 10000 images each, are used as probe sets in case of MORPH dataset. 501 face images from fa set are used as training set, while dup I (722 images) and dup II (234 images) sets are used as probe sets in case of FERET database. 735-dimensional DSAF feature vector is extracted for each difference half face image with extracted asymmetries. We use PCA [9] to reduce 735-dimensional feature vector into a 300-dimensional DSAF feature vector.

Similarly a 147-dimensional DSAF feature vector is extracted each for region 1, region 2, region 3, and region 4 of difference half face image with extracted facial asymmetries. PCA is used to reduce feature vector length to 100 in case of DSAF feature vector for each region. The number and dimensions of these patches along with DSAF feature vector dimensions are summarized in Table 6.8.

Table 6.8: Summary of DSAF features extraction

	Difference half face	Each region
Image/region size	128×128	32×64
Patch size	16×16	16×16
Number of patches	105	21
Dimension of DSAF feature vector	735	147
Dimension of DSAF feature vector after applying PCA	300	100

Rank-1 identification rates are reported in Table 6.9 for four age groups (16-20, 21-30, 31-45, and 46-60+ for MORPH database and 10-20, 21-30, 31-45, and 45-60+ for FERET database), each for difference half face, region 1, region 2, region 3 and region 4.

Table 6.9: Rank-1 recognition accuracies of different facial regions for (a) MORPH and (b) FERET aging datasets

	Rank-1 recognition accuracy for difference half face regions for older images set					Rank-1 recognition accuracy for difference half face regions for oldest images set				
Age Groups	difference half face	Region 1	Region 2	Region 3	Region 4	difference half face	Region 1	Region 2	Region 3	Region 4
16-20	57.05	46.69	52.97	52.90	**57.80**	51.10	44.40	51.00	43.35	**52.10**
21-30	52.63	45.50	53.00	54.25	**56.58**	49.15	41.00	49.20	48.60	**51.50**
31-45	**53.00**	46.85	53.35	48.95	47.50	**53.01**	39.19	47.95	46.20	45.00
46-60+	**53.30**	41.25	52.00	48.00	47.50	**46.00**	33.35	29.00	43.00	43.25

(a)

	Rank-1 recognition accuracy for difference half face regions for dup I set					Rank-1 recognition accuracy for difference half face regions for dup II set				
Age Groups	difference half face	Region 1	Region 2	Region 3	Region 4	difference half face	Region 1	Region 2	Region 3	Region 4
10-20	56.37	46.67	52.77	53.60	**58.17**	50.00	43.16	50.00	41.88	**51.70**
21-30	53.73	45.01	52.90	53.32	**55.26**	48.71	41.02	48.71	47.43	**50.00**
31-45	**52.90**	45.01	52.35	48.19	47.05	**52.13**	38.03	47.86	45.72	44.87
46-60+	**52.17**	42.25	50.00	47.82	46.95	**45.72**	32.05	24.78	41.45	41.45

(b)

On analyzing the accuracies of different age groups, it is observed that recognition accuracy decreases for older age groups, compared to younger age groups. We attribute this fact to the significant variations of facial asymmetry with increasing age. For younger age groups (16-20, 21-30 for MORPH and 10-20, 21-30 for FERET datasets), region 4 gives the highest accuracy of 57.80% and 56.58% for older image set and 52.10% and 51.50% for oldest images set on MORPH while 58.17% and 55.26% for dup I set, and 51.70% and 50.00% for dup II set for FERET datasets, respectively. This observation is attributed to the fact that this facial region contains more asymmetric information compared to other regions for younger age groups. On the other hand, for older age groups (31-45 and 46-60+), the highest recognition rates are achieved for difference half face images, both for MORPH and FERET datasets, compared to individual regional accuracies as the facial asymmetry increases with passage of time for all facial regions. Thus difference half face plays significant role in recognizing age-separated face images for older age groups while for younger age groups, region 4 is most discriminative in recognizing age-separated face images.

6.5.2 Age-Separated Face Recognition Algorithm

The objective of this component of study is to build a face recognition algorithm based on knowledge learned from role of various asymmetric facial regions in recognizing age-seperated face images. The proposed approach is formulated as follows.

- First the age group of probe face image is estimated.
- LBP [11] feature vectors are extracted for gallery and probe face images.
- DSAF feature vectors are extracted for difference half face images with extracted asymmetries, both for gallery and probe images.
- Matching scores of LBP and DSAF feature vectors are calculated for gallery and probe face images.
- To perform face recognition of age separated face images, weights are assigned to different regions of difference half face images with extracted asymmetries, based on their significance in discriminating age-seperated face images of various age groups.

The block diagram of proposed face recognition approach is shown in Figure 6.6, which is explained in the following steps.

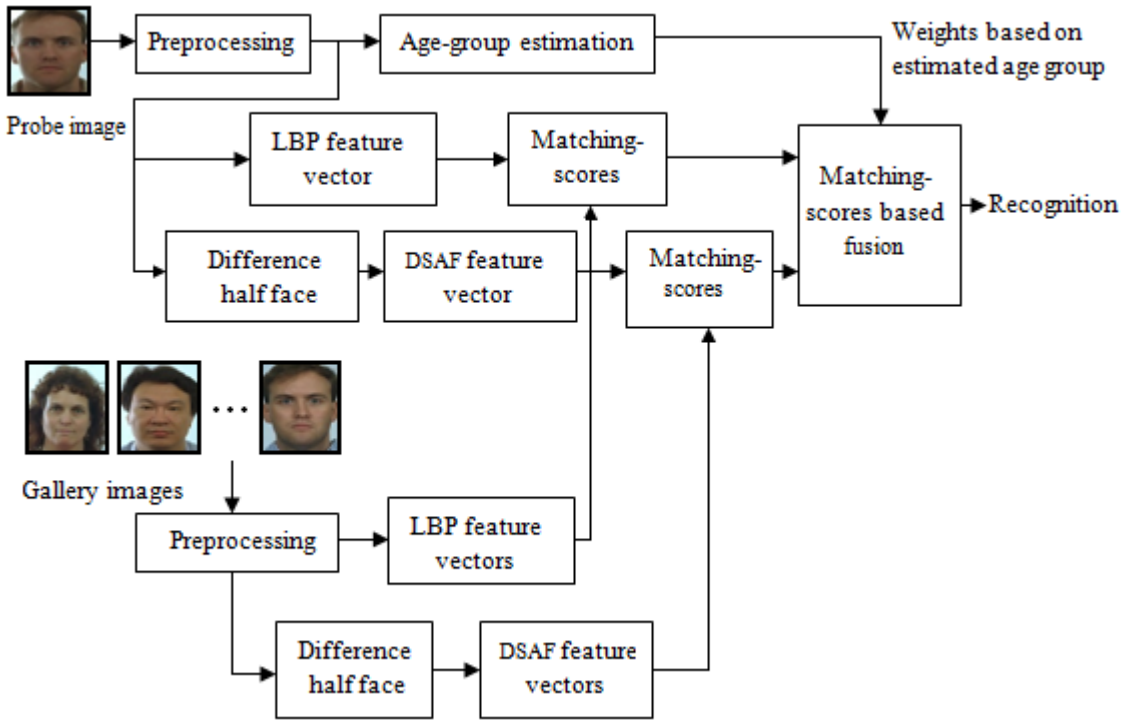


Figure 6.6: Schematic of age group estimation based face recognition algorithm

(i) To extract facial features, each face image is pre-processed.

(ii) For each face image of size 128×128 pixels, LBP features are extracted by dividing it into overlapping patches to extract local features from each patch. We divide each face image into 15×15 blocks with a patch size of 16×16 and overlapping radius of 8. We extract a 256-dimensioanl LBP featutre for each patch resulting into a 57600-dimensional LBP feature vector for each face image. PCA is used to reduce feature vector length to 300.

(iii) To extract asymmetric facial features, we use difference half face images with extracted asymmetries. A 735-dimensional DSAF feature vector is extracted for each

difference half face image and a 147-dimensional DSAF feature vector is extracted for each region of difference half face image. PCA [9] is used to reduce 735 and 147-dimensional feature vectors to 300 and 100-dimensional feature vectors, respectively.

(iv) Matching-scores of gallery and probe face images are calculated by comparing respective feature vectors of LBP and DSAF based algorithms.We get a matching-scores matrix M_j (j represents the individual algorithms) of size $A \times B$ (where A and B represent the number of probe and gallery images, respectively). The matching-scores matrix M_j has a negative polarity in our case i.e. lower matching scores represent a higher similarity and vice versa. We get six matching-scores matricesM_j, one for LBP, one for DSAF of difference half face and four for DSAF feature vectors belonging to each region.

(v) Each matching-scores matrix is normalized before fusion. We use simple *min-max* rule [31] to normalize each row (representing a probe image), on a scale of 0 to 1 as shown in equation (6.5).

$$M'_{jrow} = \frac{M_{jrow} - \min(M_{jrow})}{\max\left(M_{jrow} - \min(M_{jrow})\right) - \min(M_{jrow} - \min\left(M_{jrow}\right))} \tag{6.5}$$

Where $maxM_{jrow}$ and $minM_{jrow}$ denote the minimum and maximum values in a particular row of matching-scores matrix.

(vi) The five normalized matching score matrices, belonging to difference half face and four regions are then fused to get a combined matching score matrix M_{row} by using weighted sum rule as given in equation (6.6).

$$M_{row} = \sum_{i=1}^{5} w_i M_{rowi} \tag{6.6}$$

In equation (6.6) w_i represent weights belonging to i^{th} difference half face region and can be assigned using recognition rates of difference half face and four regions (see Table 6.9), as given in equation (6.7).

$$w_i^k = \frac{r_i^k}{\sum r_i^k} \tag{6.7}$$

Where r_i^k represent recognition rates of k^{th} age group of gallery and probe image and i^{th} difference half face or one of four regions. For example the predicted age group of a probe image is 10-20, then weights can be assigned to difference half face and each region as per the recognition accuracies, calculated for the corresponding age group, as given in Table 6.9.

The matching score matrix M_{row} is once again normalized using *min-max* rule to get M'_{row} as given in equation (6.8).

$$M'_{row} = \frac{M_{row} - \min(M_{row})}{\max(M_{row} - \min(M_{row})) - \min(M_{row} - \min(M_{row}))} \tag{6.8}$$

(vii) Finally, the matching score matrix of LBP feature vectors is fused with that obtained in equation (6.8) to get a final matching score matrix M_{sum} by using simple sum rule as given in equation (6.9).

$$M_{sum} = \sum_h M_h \tag{6.9}$$

where M_h are the matching scores belonging to LBP based matching scores and those obtained by using equation (6.9). The final matching score matrix M_{sum} is once again normalized using *min-max* rule to get M'_{sum} (see equation 6.10) which is used to calculate the rank-1 recognition rates.

$$M'_{sum} = \frac{M_{sum} - \min(M_{sum})}{\max(M_{sum} - \min(M_{sum})) - \min(M_{sum} - \min(M_{sum}))} \tag{6.10}$$

6.5.3 Evaluation and Experimental Results of Proposed Face Recognition Algorithm

Face identification experiments are performed on MORPH and FERET aging datasets as described below.

(i) In case of MORPH database, 10000 youngest face images of 10000 subjects are used as gallery set while 10000 face images each from older and oldest face image sets are used as aging probe sets, representing small and large temporal variations, respectively.

(ii) In case of FERET database, standard gallery and probe sets from FERET database are used to evaluate the performance of above proposed algorithm. 1196 face images from fa set are used as gallery set while dup I with 722 images and dup II with 234 face images are used as aging probe sets, representing small and large temporal variations respectively. As shown in Figure 6.6, the age group of probe image is predicted using the hierarchical approach proposed in section 6.2. Based on predicted age group, weights are assigned to the appropriate difference face regions depending on their significance in recognizing age-separated face images of various age groups as shown in Table 6.9 (a) and 6.9 (b) for MORPH and FERET datasets, respectively.

The Rank-1 recognition accuracies are reported for all the experiments, both for MORPH and FERET datasets, in Table 6.10.

6.5.4 Comparison with Existing Methods

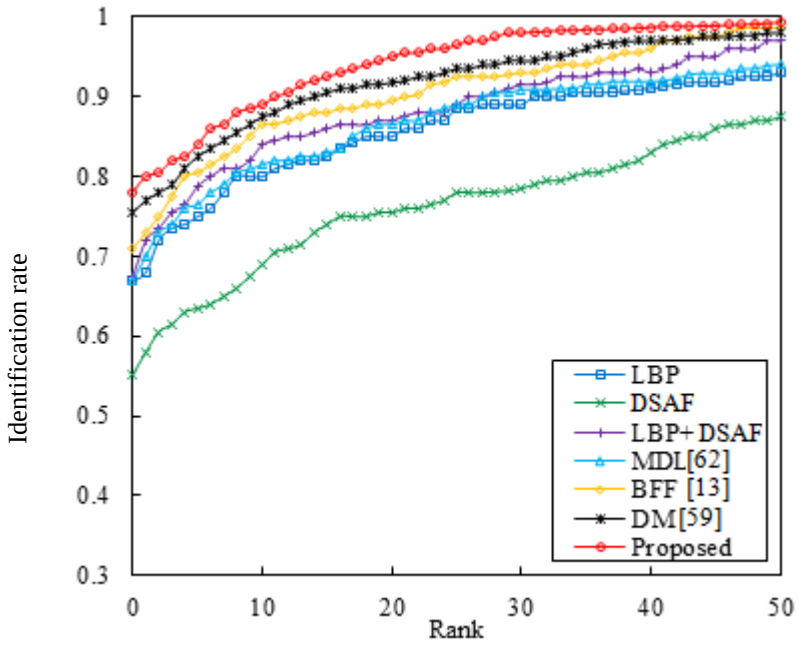
For the same experimental set up, the recognition accuracies of proposed approach are compared with following existing state-of-the-art methods.

(i) Base line algorithm, LBP [11].
(ii) Sum rule [94].
(iii) Discriminative Model (DM) [59]: In DM, MLBP and SIFT feature vectors are extracted for overlapping patches of face images. A multi-feature discriminant analysis is then performed to construct a random subspace based fusion model to recognize age-separated face images.
(iv) Bacteria foraging fusion (BFF) [13].
(v) Multi-view discriminative learning (MDL) [62].

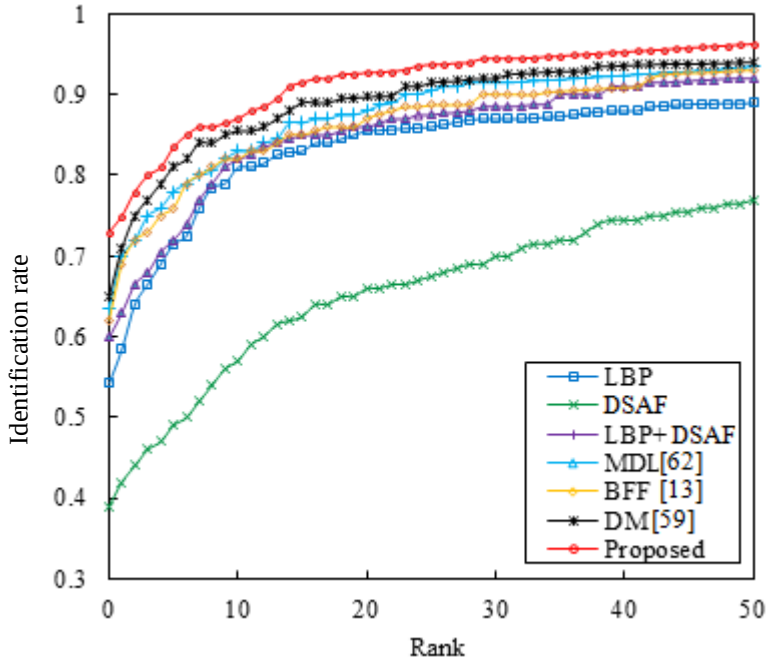
Table 6.10 Comparison of face recognition accuracies of proposed approach with existing algorithms for MORPH and FERET datasets

Approach	Facial Region (s)	Rank-1 recognition accuracy (%)			
		MORPH		FERET	
		Older images as probe set	Oldest images as probe set	Dup I as probe set	Dup II as probe set
(i) LBP [11]	Face	66.89	54.19	62.04	50.00
(ii) DSAF	Difference half face	55.20	41.00	53.04	37.17
(iii) Sum Rule (LBP+DSAF) [94]	Face, difference half face	65.54	63.52	67.45	59.82
(vi) DM [59]	Overlapping patches	75.50	65.60	69.94	64.10
(v) BFF [13]	Face, mouth, binocular, periocular	71.00	64.00	68.42	61.11
(iv) MDL [62]	Overlapping patches	67.00	59.90	67.03	60.68
(vii) Proposed approach with weights assigned to difference half face regions (actual age groups)	Face, difference half face, region 1, region 2, region 3, region 4	80.79	75.00	76.45	69.65
(viii) Proposed approach with weights assigned to difference half face regions (estimated age groups)	Face, difference half face, region 1, region 2, region 3, region 4	**78.10**	**73.01**	**73.82**	**67.09**

The CMC curves comparing the proposed approach with existing algorithms are shown in Figures 6.7(a) and 6.7 (b) for MORPH datasets and Figure 6.8 (a) and 6.8(b) for FERET datasets, respectively.

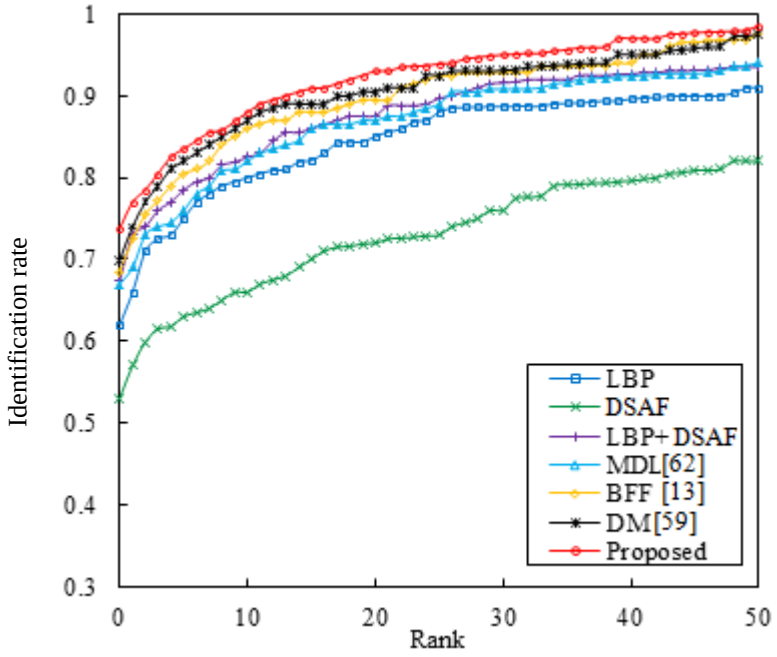


(a)

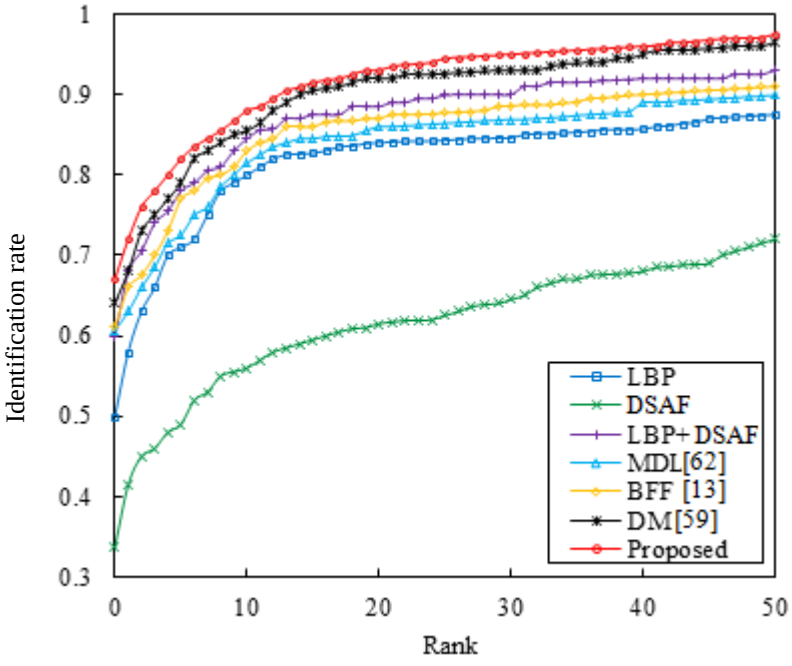


(b)

Figure 6.7: CMC curves showing face identification performance for (a) small and (b) large temporal variations on MORPH datasets



(a)



(b)

Figure 6.8: CMC curves showing face identification performance for (a) small and (b) large temporal variations on FERET datasets

6.5.5 Computational Complexity Analysis

It is worth underlining the two-phased nature of proposed age group estimation based face recognition method. The first phase contains the procedures for age group estimation. Here, the computational complexity of dimension selection through F-test is $O(n^2)$. The computational complexity of calculating discriminating scores using CCA is $O(nd^2) + O(d^3)$ where $d = max(D_x, D_y)$. For classification stage, since we only need to train $N - 1$ binary SVM classifiers, hence the order of complexity is $O(\log N)$.

The second phase consists of recognition of age-separated face images. Here, For LBP feature extraction, the order of complexity is $O(N_p)$, where N_p is number of pixels in face image. Extraction of AFF feature vectors results into linear complexity of $O(GN_{dp} + 4LN_{dp})$ for APs, while $O(N_{dp}^2)$ for Hu moment invariants, where *G*, *L*, and N_{dp} represent number of gray levels in difference half face image, number of transformation levels in APs and number of pixels in difference half face image, respectively.

The above analysis shows that the calculation of discriminating scores using CCA and extraction of Hu moment invariants are computationally most demanding stages in age group estimation and face recognition components of proposed approach, respectively.

6.6 Results Related Discussion

The research conducted in this study covers two important facial aging aspects: age group estimation, and recognition of age-separated face images. For the sake of clarity, the performance of these two aspects is discussed separately as follows.

6.6.1 Performance of Age Group Estimation Algorithm:

(i) The proposed method achieved an overall age group estimation accuracy of 93.86% and 92.98% on MORPH and FERET datasets compared to 79.41% and 77.99% given by the LBP variants proposed in [49], while Face++ [25] gave an overall accuracy of 87.26 and 86.23 % on MORPH and FERET datasets respectively. Hence proposed method

outperformed both the existing methods considerably. We attribute this outperformance to the invariance of proposed method to extrinsic facial variations like facial make up. Hence we conclude that facial asymmetry is a strong indicator of age and may effectively be used in age group estimation task.

(ii) Due to insensitivity of facial asymmetry to extrinsic variations, the older age groups are least confused with corresponding younger age groups, which remains a challenge for previous related studies [28], [50] where older age groups are more vulnerable to be confused with younger ones due to use of facial make up by certain age groups.

6.6.2 Performance of Face Recognition Algorithm

(i) In case of MORPH database, the Rank-1 recognition accuracies of 78.10% and 73.01% are achieved for older and oldest image sets, respectively. In case of FERET database, the Rank-1 recognition accuracies of 73.82% and 67.09% are achieved for dup I and dup II sets representing small and large temporal variations, respectively.

(ii) In all face identification experiments the proposed approach outperforms the sum rule fusion applied on face and difference half face and three existing state-of-the-art approaches, DM [59], MDL [62], and BFF [13], on both the MORPH and FERET datasets, for small and large temporal variations.

(iii) To analyze the errors introduced by age group estimation of probe image, we computed the rank-1 accuracy of face, difference half face, region 1, region 2, region 3 and region 4, for both scenarios: (i) with actual age groups and (ii) with estimated age groups. We observed a decrease in recognition accuracy of 3.32% and 2.65% for older and oldest image sets in case of MORPH datasets. In case of FERET, accuracy decreases by 3.44% and 3.67 % for dup I and dup II sets respectively. The minimal decrease in accuracy both for MORPH and FERET aging datasets show the superior age group estimation performance of proposed approach.

The Rank-1 identification results across small and large temporal variations suggest that integrating the existing facial feature technique (LBP) with asymmetric facial features

(DSAF) along with the knowledge learned from relevant accuracy of certain asymmetric facial regions can significantly boost the recognition accuracy of age separated face images.

6.7 Summary

In this chapter we have explored the role of facial asymmetry in age group estimation and face recognition of age-separated face images. The relevance of various asymmetric facial regions is used to improve face recognition across small and large temporal variations. As a facial feature, facial asymmetry has been used along with existing feature extraction technique, the local binary patterns (LBP), to improve the accuracy of face recognition of age-separated face images. The proposed method clearly shows an improvement in face recognition accuracy when knowledge learned from age group estimation is integrated with face recognition algorithm.

Next we plan to use some of these insights to develop a framework capable of estimating multiple demographic estimates of a given query face image before recognizing it across temporal variations. Integration of the knowledge learned from multiple demographic estimates into a face recognition algorithm will enhance the recognition performance of age-separated face images, considerably.

Chapter 7

DEMOGRAPHIC ESTIMATION BASED RECOGNITION

The introduced Densely Sampled Asymmetric Facial features (DSAF) in Chapter 4, provide a strong base to estimate age group and recognize age separated face images. In order to improve recognition performance of age-separated face images, knowledge learned from other demographic estimates such as gender and race may also be integrated with face recognition algorithm. In this chapter, therefore, we learn facial asymmetry based estimates of age group, gender, and race of a query face image and integrate this knowledge with face recognition algorithm to enhance the recognition performance of age-separated face images.

7.1 Introduction

Recognition of face images and demographic estimation are important yet challenging problems. These challenges mainly include pose, illumination, expression, and aging variations. Among these variations, aging variation is still a complex face recognition problem and hence receiving continuous attention of research community. Developing algorithms for recognition of age-separated face images finds many applications such as, immigration documents and security services [38].

Facial appearance varies with increasing age due to numerous factors such as gender, race, nutrition, and health [95]. Bilateral facial asymmetry, which refers to non-correspondence in shape, size, and location of left and right-sided facial landmarks [82], is another important factor, which increases with age [96]. During formative years, human face is more symmetric while in the later years, facial asymmetry is more prominent.

In this chapter, an integrated approach for building an age-invariant face recognition system is presented. The proposed approach involves (i) facial asymmetry based demographic estimation and (ii) subsequent recognition of age-separated face images. Densely sampled asymmetric features (DSAF) are extracted from a face image and a feature selection method will be used to estimate age group, gender, and race of a query face image. In contrast to previous chapter where only the age group estimation based knowledge was incorporated into face recognition algorithm, the current chapter focuses on integration of age group, gender, and race estimates into a face recognition algorithm to enhance recognition performance of age-separated face images.

7.2 Demographic Estimation

The proposed approach for demographic estimation (age group, gender, and race) consists of three main components: (i) pre-processing, (ii) feature representation with asymmetric facial features, and (iii) demographic estimation using SVMs as classifiers.

7.2.1 Pre-Processing

For proposed demographic estimation task, standard MORPH and FERET facial aging databases are used. To compensate for facial variations, a face pre-processing method is devised, as described below.

(i) First of all, color face images are converted into gray-scale images to correct for unwanted color cast. For this purpose a luminance model adopted by NTSC and JPEG as is given in Chapter 6.

(ii) To mitigate the effect of in-plane rotation and translation, face images are aligned based on fixed eye locations, such that images are upright and cropped into 128×128 pixels size with equal inter-pupillary distance. Eye locations are detected using a publically available state-of-the-art method, Face++ [25].

(iii) Finally, to correct illumination variations due to shadow and underexposure, we use widely used Difference of Gaussian (DoG) filtering as suggested in [97]. Figure 7.1

shows face pre-processing results for two age-separated face images of one subject from MORPH database.

(a) (b) (c) (d)

Figure 7.1: Face pre-processing of two age-separated face images of same the subject from MORPH database: (a) input images (b) Gray-scale images (c) images aligned based on fixed eye locations (d) illumination correction

7.2.2 Feature Selection and Classification

Since we aim to use introduced DSAF facial features for demographic estimation (age group, gender, and race), it is necessary to have a feature selection method for this purpose. To perform demographic estimation task, feature selection is performed by using AdaBoost learning algorithm [98]. AdaBoost employs a combination of weak classifiers to determine a subset of discriminating features for a particular classification task, resulting into enhanced classification rates. In this study, the dimensionality of DSAF feature vector is 735. We aim at selecting a subset of DSAF feature vector each for age group, gender, and race estimation tasks, using AdaBoost. In proposed work, three SVMs are trained as weak classifiers for feature selection, with RBF as kernel. A multiclass SVM is used for age group estimation task, while two binary SVM classifiers are used each for the gender and race classification tasks.

The boosting procedure to select a discriminating feature subset for age group, gender and race estimation task is given in Algorithm 6.1. In all of our experiments we use a 140-dimensional feature vector, learned through AdaBoost algorithm each for age group, gender, and race estimation tasks. The labels depend upon different demographic

estimation tasks. For n number of age groups, we use labels $(0, 1, 2, \dots, n)$, while for gender and race estimation, binary labels $(0, 1)$ are used to represent (male, female) or (white, others), respectively.

Given training images $T = (a_1, b_1), \dots, (a_n, b_n)$, where $a_i \in A$ and $b_i \in B = \{0, 1, 2, 3\}$ for age group, $\{0, 1\}$ for the gender and race estimation tasks.

Initialize weights for training samples $w_i^{(1)} = \frac{1}{n}, \Re = \Phi$, where $n = 0, 1, 2,$

For $t = 1, \dots T$

Normalize the weights $w_i^t = \frac{w_i^t}{\sum_{i=t}^{n} w_i^t}$

Use SVM with RBF kernel to train a weak classifier h_t and select training sample subset s_t, such that $s_t \subseteq T$

Compute the weighted error of weak classifier h_r

$$\epsilon_t = \sum_{i=1}^{n} w_t(i), b_i \neq h_r(a_i)$$

Find the discriminating features index: $\hat{r} = \operatorname{argmin} e_t$

Update the selected feature indexes: $\Re = \Re \cup \{\hat{r}\}$

Update the weights of training samples:

$$w_{t+1,i} = w_t^i \beta_t^{1-\epsilon h_r^{\wedge}(a_i, b_i)}$$

where $\beta_t = \frac{1}{2} \ln(\frac{1-e_{\hat{r}}}{e_{\hat{r}}})$

Result: desired feature set $F = \{a_i(\Re)\}_{i=1}^{n}$

Algorithm 7.1: The boosting algorithm for demographic informative feature selection

7.2.3 Experimental Results of Demographic Estimation Algorithm

We perform age group, gender and race classification on difference half face images with extracted facial asymmetries from MORPH and FERET databases using a five-fold cross

validation protocol. Due to significantly imbalanced race distribution in MORPH and FERET databases, we perform binary race estimation between white and other races (Black, Hispanic, Asian, and African-American). For each classification task, the results are reported in the form of confusion matrix (see Table 7.1) with experimental protocol for each database described below.

Experiments on MORPH database: A subset consisting of 10000 frontal face images from MORPH database is selected. 5000 face images are used in training while 5000 face images are used in test sets with age ranges of 16-20, 21-30, 31-45, and 46-60+. The performance of proposed method is compared with Face++ [25] for age group, gender, and race estimation on MORPH database as shown in Table 7.1 (a). The proposed method achieves an overall age group estimation accuracy of 94.54% with a standard deviation of 1.3 across five folds, compared to the state-of-the-art method Face++, which achieves an overall accuracy of 89.15% with a standard deviation of 1.5 across five folds. In case of Face++, the subjects (especially for the age range of 46-60+) are found to be more baffled with subjects in younger age groups, mainly because of use of facial make-up, resulting into younger appearance of face images. On the other hand, asymmetric facial features, used in proposed method, are difficult to manipulate by using facial make-up and hence the subjects are less confused with corresponding younger age groups.

The results of gender estimation performance of proposed approach and Face++ are shown in Table 7.1 (b). The proposed method achieves an overall accuracy of 83.19% with a standard deviation of 0.5, compared to Face++ which achieves an overall accuracy of 77.86% with a standard deviation of 0.4. An interesting observation is higher misclassification rates for females than males achieved by the proposed approach. One possible explanation for this higher misclassification is frequent extrinsic variations of female faces, such as the facial make-up and varying eyebrow styles.

Table 7.1: Confusion matrices showing classification accuracies of proposed and Face++ methods for (a) age group, (b) gender, and (c) race estimation tasks on MORPH datasets

Training age groups (# subjects)	Predicted age groups							
	Proposed				Face++ [25]			
	16-20	21-30	31-45	46-60+	16-20	21-30	31-45	46-60+
16-20	**92.45**	6.08	1.21	0.24	88.80	7.54	2.43	1.21
21-30	1.47	**95.21**	2.60	0.70	3.59	92.04	2.32	2.04
31-45	0.18	1.40	**96.10**	2.29	0.75	3.93	93.53	1.78
46-60+	1.6	1.12	2.88	**94.4**	3.20	5.28	9.28	82.24
Overall accuracy	**94.54 ±1.3**				89.15 ±1.5			

(a)

Training gender (# subjects)	Predicted gender			
	Proposed		Face++ [25]	
	Male	Female	Male	Female
Male	**85.94**	14.06	80.70	19.30
Female	19.55	**80.45**	24.98	75.02
Overall accuracy	**83.19 ±0.5**		77.86 ±0.4	

(b)

Training race (# subjects)	Predicted race			
	Proposed		Face++ [25]	
	White	Others	White	Others
White	**90.00**	10.00	83.02	16.98
Others	29.82	**70.18**	32.00	68.00
Overall accuracy	**80.09 ±1.2**		75.51 ±1.1	

(c)

The race estimation experiments conducted on MORPH database classify subjects between white and other subjects. The race estimation results are shown in Table 7.1 (c) depicting the superior of proposed approach compared to Face++. The proposed method achieves an overall accuracy of 80.09 with a standard deviation of 1.2 compared to Face++ which achieves an overall accuracy of 75.51% with a standard deviation of 1.1. It can be seen that proposed approach is better at estimating the white race compared to other races.

Experiments on FERET database: 1196 frontal face images from fa set of FERET database are selected. 598 images are used in training while 598 images are used in testing for demographic estimation tasks. The performance of proposed and Face++ methods for age group, gender and race estimation on FERET database is shown in Table 7.2.

Table 7.2: Confusion matrices showing the classification accuracies of proposed and Face++ methods for (a) age group, (b) gender, and (c) race estimation tasks on FERET datasets

Training age groups	Predicted age groups							
	Proposed				Face++ [25]			
	10-20	21-30	31-45	46-60+	10-20	21-30	31-45	46-60+
10-20	**85.04**	6.41	4.70	3.84	83.33	8.54	6.40	1.73
21-30	3.07	**84.61**	6.15	6.15	3.84	83.84	12.32	0
31-45	0	8.33	**79.76**	11.90	3.57	3.57	80.94	11.90
46-60+	0	5.33	12.00	**82.66**	0	8.67	13.33	78.00
Overall accuracy	**83.01 ±0.9**				81.52 ±1.1			

(a)

Training gender	Predicted gender			
	Proposed		Face++ [25]	
	Male	Female	Male	Female
Male	**82.89**	17.11	78.94	21.06
Female	18.06	**81.94**	26.39	73.61
Overall accuracy	**82.42 ±0.7**		76.28 ±0.5	

(b)

Training race	Predicted race			
	Proposed		Face++ [25]	
	White	Others	White	Others
White	**84.44**	15.56	82.00	18.00
Others	27.40	**72.60**	32.88	67.12
Overall accuracy	**78.52 ±1.0**		74.56 ±0.9	

(c)

The proposed method achieves an overall age group estimation accuracy of 83.01% with a standard deviation of 0.9 across a five-fold cross validation. On the other hand Face++ achieves an overall age group accuracy of 81.52% for the same experimental protocol.

For gender estimation task, the proposed method achieves an overall estimation accuracy of 82.42% with a standard deviation of 0.7 compared to 76.28% achieved by Face++. On

comparison, it is observed that misclassification rate of female subjects is greater than male subjects in case of FERET database, similar to trend observed for subjects in MORPH database. The proposed approach achieves an overall race estimation accuracy of 78.52% with a standard deviation of 1 across five folds compared to the race estimation accuracy of 76.28% achieved by Face++, for the same experimental protocol. The proposed approach gives a higher misclassification rate for white race compared to other races (27.40% vs. 15.56%), which is attributed to large within group variety of other races.

7.3 Recognition of Age-Separated Face Images

The second part of this study consists of recognition of age-separated face images based on knowledge learned from demographic estimates (age group, gender, and race). The proposed approach to incorporate the demographic information is described in the following steps.

(i) Estimate the age group, gender and race of a given probe image.

(ii) Extract densely sampled asymmetric features (DSAF) for different asymmetric facial regions.

(iii) To match given probe image with one in gallery, assign weights to different asymmetric facial regions based on estimated age group, gender, and race along with significance of each asymmetric facial region in recognizing age-separated face images, as illustrated in following subsection.

7.3.1 Significance of Different Asymmetric Facial Regions in Recognizing Age-Separated Face Images

To understand the role of different asymmetric facial regions in recognizing age-separated face images belonging to three different demographic attributes (age group, gender, and race), each 128×64 difference half face image with extracted asymmetries is divided into four regions region 1, region 2, region 3, and region 4, such that the size of each region is 32×64 pixels. Each region is further divided into 21 overlapping patches of size 16×16 pixels with overlapping radius of 8 pixels. A 7-dimensional Hu moment

invariants feature vector is extracted for each patch resulting into 21×7=147-dimensional DSAF feature vector each for region. DSAF feature vectors for each for difference half face image with extracted asymmetries and four regions are used for recognition task with Rank-1 recognition accuracies for age group, gender and race classification tasks reported in Table 7.3.

Table 7.3: Rank-1 identification accuracies for different asymmetric facial regions for age group, gender, and race classification on (a) MORPH, and (b) FERET datasets

		Rank-1 recognition accuracy for difference half face regions for older images set					Rank-1 recognition accuracy for difference half face regions for oldest images set				
		Difference half face	Region 1	Region 2	Region 3	Region 4	Difference half face	Region 1	Region 2	Region 3	Region 4
Age groups	16-20	56.81	45.95	52.55	51.94	**56.93**	50.48	45.01	51.00	51.94	**53.04**
	21-30	50.98	45.00	52.95	53.94	**55.00**	50.00	43.94	48.02	48.02	**53.02**
	31-45	**54.00**	48.00	53.96	48.95	45.99	**53.02**	37.97	45.00	47.11	46.13
	46-60+	**55.04**	42.72	51.04	48.96	49.92	**48.00**	39.20	33.92	45.92	43.04
Gender	Male	**59.99**	55.00	51.95	50.05	50.05	**52.99**	48.00	44.02	43.29	48.49
	Female	54.00	**56.08**	50.02	51.01	54.00	46.85	42.59	47.32	47.32	**51.11**
Race	White	62.79	**64.00**	61.06	60.00	60.00	56.00	**56.97**	54.88	48.83	51.16
	Others	**56.00**	53.15	52.00	52.00	50.00	**50.00**	47.19	44.00	44.00	42.00

(a)

		Rank-1 recognition accuracy for difference half face regions for dup I set					Rank-1 recognition accuracy for difference half face regions for dup II set				
		Difference half face	Region 1	Region 2	Region 3	Region 4	Difference half face	Region 1	Region 2	Region 3	Region 4
Age groups	10-20	55.55	50.00	53.41	50.00	**58.11**	50.00	49.57	51.28	51.28	**53.84**
	21-30	45.38	45.38	47.69	49.23	**50.00**	48.46	50.00	47.69	45.38	**52.30**
	31-45	**52.38**	45.23	47.61	46.42	47.61	**53.57**	50.00	50.00	51.19	52.38
	46-60+	**54.76**	50.00	50.00	45.33	45.33	**53.33**	52.00	48.00	50.00	48.00
Gender	Male	**60.00**	58.00	57.89	55.00	55.00	**58.00**	55.00	55.00	52.10	50.00
	Female	57.87	**61.57**	55.55	55.55	57.87	60.18	**62.03**	57.87	57.87	55.55
Race	White	62.00	**65.55**	58.00	50.00	50.00	58.00	**62.00**	50.00	50.00	50.00
	Others	**60.27**	58.21	58.21	50.00	45.89	**58.21**	50.00	43.15	45.20	50.68

(b)

We use 10000 youngest face images as training set, while older and oldest face image sets (with 10000 images each) are used as probe sets in case of MORPH dataset. In case of FERET database, we use 501 face images from fa set as training set, while dup I (722 images) and dup II (234 images) sets are used as probe sets.

By comparing the accuracies of different age groups, it is observed that recognition accuracy decreases for older age groups, compared to younger age groups for age group classification task. We attribute this fact to the significant variations of facial asymmetry with increasing age. For younger age groups (16-20, 21-30 for MORPH and 10-20, 21-30 for FERET datasets), region 4 gives the highest accuracy of 56.93% and 55% for older image set and 53.04% and 53.02% for oldest images set on MORPH while 58.11% and 50.00% for dup I set, and 53.84% and 52.30% for dup II set for FERET datasets, respectively. We attribute this observation to the fact that this facial region contains more asymmetric information, compared to other regions for younger age groups. On the other hand, for older age groups (31-45 and 46-60+), we achieve the highest recognition rates for difference half face images, both for MORPH and FERET data sets compared to individual accuracies for all four facial regions. Thus difference half face plays significant role in recognizing age-separated face images for older age groups while for younger age groups, region 4 is the most discriminative in recognizing age-separated face images. It can be observed that we get slightly different recognition accuracies for different age groups for similar face images datasets used in current and previous study (presented in Chapter 6) due to difference in pre-processing schemes employed in both approaches.

In case of gender classification, difference half face image with extracted asymmetries gives the highest recognition accuracy for male subjects while region 1 gives the highest recognition accuracy for female subjects, both for MORPH and FERET databases. Finally for race classification, region 1 gives the highest recognition accuracy for white subjects while difference half face image gives the highest recognition accuracy for subjects belonging to other races, both for FERET and MORPH databases.

The results presented in Table 7.3 suggest that individual recognition accuracies of difference half face images with extracted asymmetries and four regions for age group, gender, and race classification can be incorporated with face recognition algorithm to improve face recognition performance, as described in the following section.

7.3.2 Face Recognition Algorithm Inspired from Demographic Estimates

An important component of this study is to demonstrate that the knowledge learned from demographic estimation of age group, gender, and race can be used to enhance the performance of face recognition algorithm. The proposed approach to incorporate such knowledge mainly consists of following three steps.

(i) Estimate the age group, gender, and race for query face image.
(ii) Extract densely sampled asymmetric features (DSAF) for different asymmetric facial regions (region 1, region 2, region 3, and region 4).
(iii) To match a given probe image with a gallery image, assign weights to different asymmetric facial regions based on estimated age group, gender and race and relevance of a specific asymmetric facial region using recognition accuracies presented in Table 7.3.

Such a proposed approach is termed as *demographic estimation based fusion scheme (DEFS)* for recognition of age-separated face images as illustrated in the following subsection.

7.3.3 Demographic Estimation based Fusion Scheme (DEFS) for Recognition of Age-Separated Face Images

The block diagram of proposed DEFS face recognition algorithm is shown in Figure 7.2.

The proposed demographic estimation based fusion scheme (DEFS) consists of two main parts:

(i) Demographic estimation: The objective of this part is to estimate age group, gender, and race of a query face image using the algorithm described in section 7.2.

(ii) Face matching: After estimating age group, gender, and race of query face image, the next step is to match its identity with one of gallery images through a matching algorithm as explained in the following steps.

(1) DSAF feature extraction: DSAF features are extracted for each difference half face and four regions. DSAF features of probe face image are matched with corresponding gallery features to obtain matching scores of difference half face image and four regions.

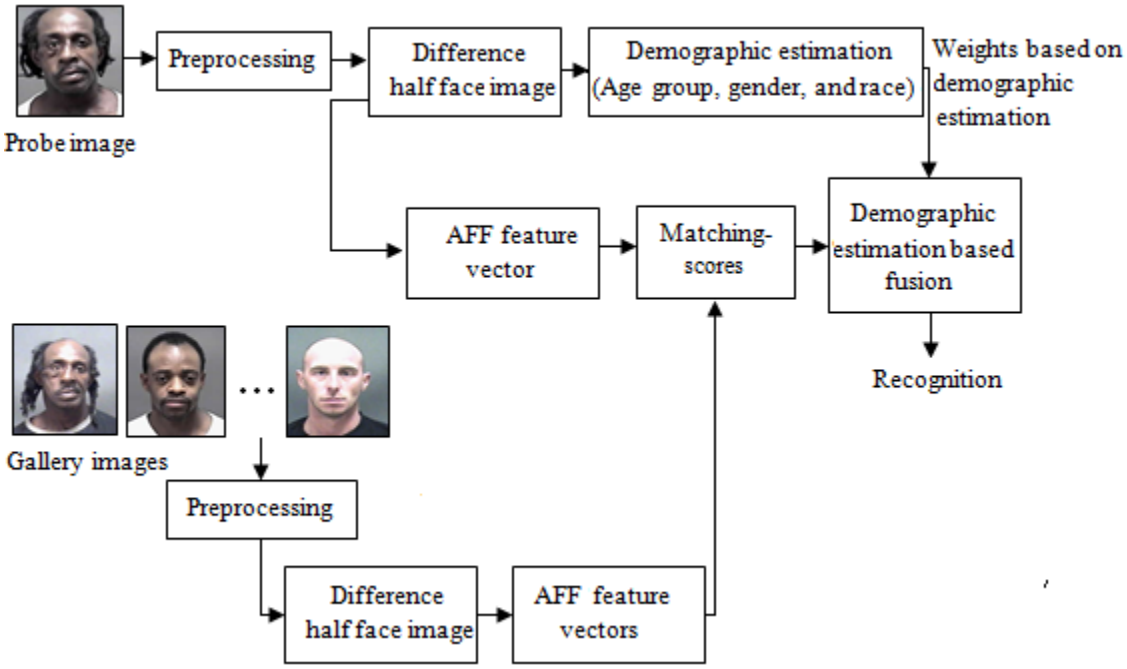


Figure 7.2: Proposed demographic estimation based fusion scheme (DEFS) for recognition of age-separated face images

(2) Weighted score-level fusion: Matching scores pertaining to four regions (region 1, region 2, region 3, and region 4) are combined using weighted sum rule [95] as given in equation (7.1).

$$S_{combined} = \sum_{i,j} w_i^j s_i^j \tag{7.1}$$

where w_i^j represents the weight pertaining to i^{th} region and j^{th} demographic estimate and s_i^j represents the corresponding matching score. The weights in proposed DEFS

approach can be learned using the individual recognition accuracies of region 1, region 2, region 3, and region 4, for three demographic estimates (i.e., age group, gender, and race), using Table 7.3. Equation (7.2) illustrates the calculation of weights for proposed fusion approach.

$$w_i^j = \frac{r_i^j}{\sum r_i^j} \tag{7.2}$$

where r_i^j represents the recognition accuracy of i^{th} asymmetric facial region and j^{th} demographic estimate, such that i={region 1, region 2, region 3, region 4}, and j={age group 1, age group 2, age group 3, age group 4}, j={male, female}, and j={white, others}, for age group, gender, and race estimates, respectively. The combined matching scores are finally used to recognize the probe image against gallery images.

7.4 Experiments and Results of Proposed Face Recognition Approach

The performance of proposed DEFS algorithm is evaluated on two standard facial aging databases: the MORPH and FERET as described below.

7.4.1 Experiments on MORPH Database

In case of MORPH database, 10000 youngest face images of 10000 subjects are used as gallery set while 10000 face images each from older and oldest face image sets are used as aging probe sets, representing small and large temporal variations, respectively.

The Rank-1 recognition accuracies are reported both for older and oldest face images sets as shown in Table 7.4. The Rank-1 recognition accuracies of 80.05% and 75.00% for MORPH for older and oldest image sets are achieved for proposed DEFS approach.

7.4.2 Experiments on FERET Database

In case of FERET database, standard gallery and probe sets from FERET database are used to evaluate the performance of above described algorithm. 1196 face images from fa set are used as gallery set while dup I with 722 images and dup II with 234 face images are used as aging probe sets, representing small and large temporal variations

respectively. The Rank-1 recognition accuracies of 77% and 75.21% for dup I and dup II sets are achieved for FERET database as shown in Table 7.4.

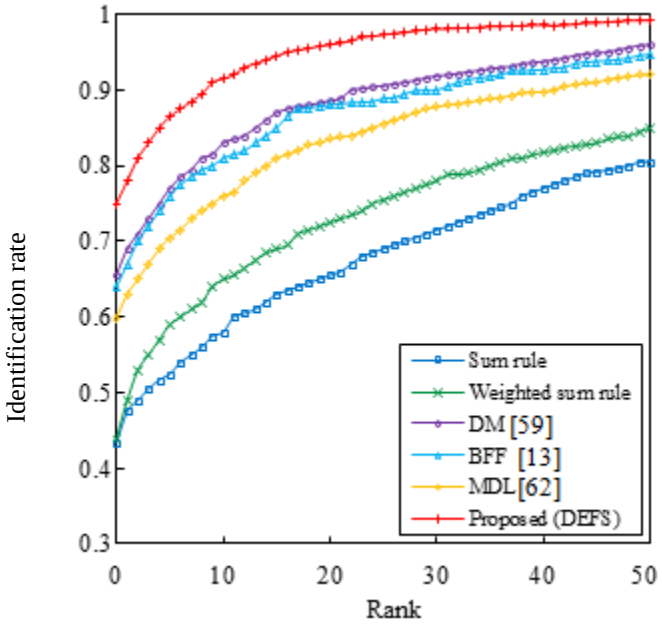
7.4.3 Comparison with Existing Methods

For the same experimental set up, the recognition accuracies of proposed DEFS approach are compared with sum rule and weighted sum rule [95] in addition to following state-of-the-art methods. (i) Discriminative Model (DM) [59]. (ii) Bacteria foraging fusion (BFF) [13]. (iii) Multi-view discriminative learning (MDL) [62].

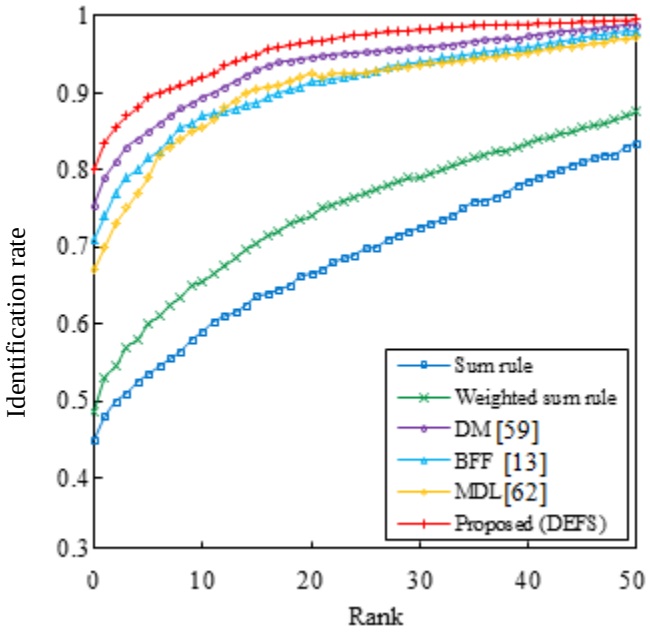
Table 7.4: Rank-1 recognition accuracies of proposed and existing methods

Approach	Facial Region (s)	Rank-1 recognition accuracy (%)			
		MORPH		FERET	
		Older images as probe set	Oldest images as probe set	Dup I as probe set	Dup II as probe set
(i) Sum Rule [95]	Region 1, region 2, region 3, region 4	45.00	43.40	44.04	42.30
(ii) Weighted sum rule [95]	Region 1, region 2, region 3, region 4	48.71	44.02	46.53	44.87
(iii) DM [59]	Overlapping patches	75.50	65.60	69.94	64.10
(iv) BFF [13]	Face, mouth, binocular, periocular	71.00	64.00	68.42	61.11
(v) MDL [62]	Overlapping patches	67.00	59.90	67.03	60.68
(vi) Proposed DEFS approach with weights assigned to difference half face regions (actual age groups, gender and race)	difference half face, region 1, region 2, region 3, region 4	82.00	78.61	80.05	77.77
(vii) Proposed DEFS approach with weights assigned to difference half face regions (estimated age groups, gender, and race)	difference half face, region 1, region 2, region 3, region 4	80.05	75.00	77.00	75.21

The CMC curves comparing the proposed approach with above mentioned methods are shown in Figure 6.3 for MORPH and Figure 6.4 for FERET aging datasets respectively.

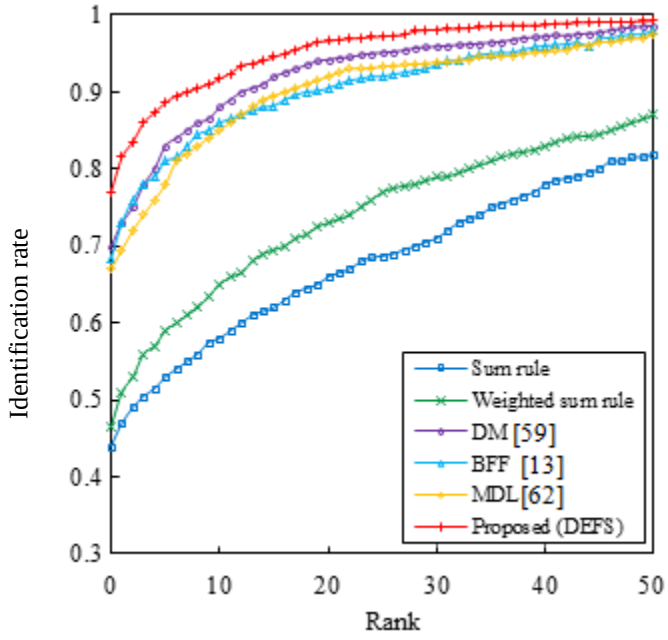


(a)

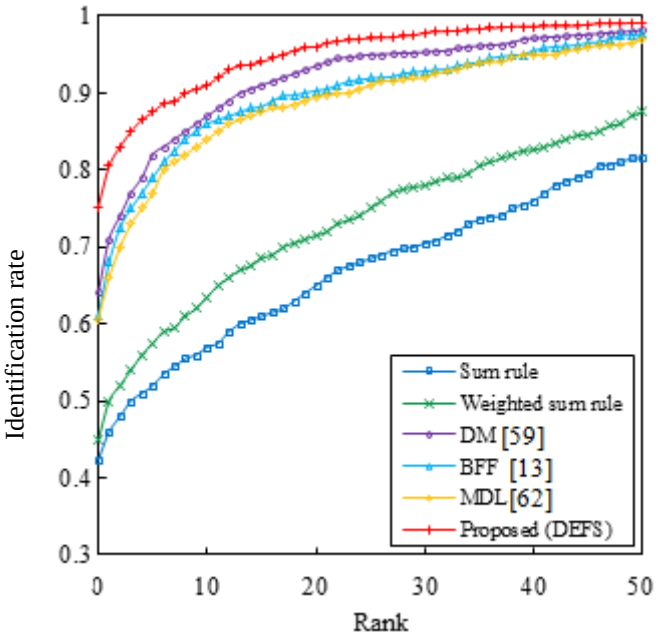


(b)

Figure 7.3: CMC curves showing the Rank-1 identification performance of proposed and existing approaches on (a) older, and (b) oldest image sets for MORPH database



(a)



(b)

Figure 7.4: CMC curves showing the Rank-1 identification performance of proposed and existing approaches on (a) dup I, and (b) dup II image sets for FERET database

7.4.4 Computational Complexity Analysis

In the two-phased proposed demographic estimation based face recognition method, the first phase contains the procedures for demographic estimation. Here, the computational complexity of Adaboost learning is $O(RNK)$, where R represents the number of rounds, N represents the number of face images with K features each.

The second phase consists of recognition of age-separated face images. Here, extraction of DSAF feature vectors results into linear complexity of $O(GN_{dp} + 4LN_{dp})$ for APs, while $O(N_{dp}^2)$ for Hu moment invariants, where G, L, and N_{dp} represent number of gray levels in difference half face image, number of transformation levels in APs and number of pixels in difference half face image, respectively.

The above analysis shows that the calculation of Hu moment invariants are computationally most demanding stage in face recognition component of proposed approach.

7.5 Results Related Discussion

For the sake of clarity, the performance of proposed face recognition algorithm is discussed separately, both for MORPH and FERET databases.

7.5.1 Results on MORPH Database

For MORPH older and oldest aging datasets, the proposed DEFS approach outperforms the existing state-of-the-art methods, DM [59], BFF [13], and MDL [62], when probe images belong to older and oldest image sets of MORPH database. The proposed approach achieves the Rank-1 identification accuracies of 80.05% and 75% for older and oldest face image sets. The experiments on older and oldest face images sets are useful in analyzing the effect of increasing age on face recognition performance. The proposed approach also achieves higher recognition accuracy compared to traditional fusion schemes such as sum rule and weighted sum rule [94]. The recognition accuracy of simple sum rule achieved by fusing the matching scores of region 1, region 2, region 3,

and region 4 of difference half face image with extracted asymmetries is 45 % and 43.40% for older and oldest face image sets which is far less than the corresponding recognition accuracies achieved by proposed approach. This fact is attributed to the incorporation of age group, gender and, race estimation into face recognition algorithm.

7.5.2 Results on FERET Database

Similar to the results on MORPH database, the recognition accuracies achieved on FERET aging datasets (dup I and dup II) suggest that the proposed DEFS approach outperforms the existing state-of-the-art methods DM [59], BFF [13], and, MDL [62], in addition to traditional fusion schemes, the sum rule and weighted sum rule [94]. The proposed approach achieves a Rank-1 recognition accuracy of 77% for dup I representing small temporal variations. For the experiment, where images with large temporal variations (dup II set) are used as probe set, the proposed approach achieves a Rank-1 recognition accuracy of 75.21%. This is lower than dup I accuracy due to change in facial asymmetry with increasing age.

7.5.3 Errors Produced by Demographic Estimation in recognizing Age-separated Face Images

In order to analyze the errors introduced by demographic estimation of age group, gender, and race in proposed DEFS approach, we have also presented the recognition accuracies for the scenario if actual age group, gender, and race information (provided by ground truth information by respective databases) is used, instead of respective estimates (see row vi, Table 7.4). It can be seen that better the demographic estimation of a given probe image, lesser is the error introduced in recognizing a given age-separated probe face image and vice-versa. In the current study, the proposed approach induces an error less than 5% for demographic estimation of age group gender and race for all face recognition experiments (see row vi and vii, Table 7.4).

7.6 Summary

Human face contains a number of age-specific features. Facial asymmetry is one of such intrinsic facial features, which is a strong indicator of age group, gender, and race. The study presented in this chapter presents role of facial asymmetry in demographic estimation and its integration into face recognition algorithm to enhance the performance. The effect of different asymmetric facial regions in recognizing age-separated face images is presented. The proposed DEFS approach suggests that incorporating the knowledge learned from demographic estimation of age group, gender, and race into face recognition algorithm leads to enhanced recognition performance. The proposed approach demonstrates its effectiveness in recognizing age-separated face images on two widely used facial aging databases compared to some existing state-of-the-art approaches.

Chapter 8

CONCLUSIONS AND FUTURE DIRECTIONS

This dissertation studied the challenging problem of facial asymmetry evaluation, its role in demographic estimation and recognition of age-separated face images. The primary contributions were realized by extracting asymmetric facial features for demographic estimation and recognition of age-separated face images. Each study has made contribution to improve the recognition performance of age-separated face images. The results of these studies are an improvement to existing state-of-the art approaches that are of interest in public security, criminal investigations, and law enforcement applications.

8.1 Contributions

In chapter 2, a brief but wide ranging literature review was presented with pre-requisites of modern face recognition systems.

In Chapter 3, we presented a detailed overview of existing feature representation and feature extraction methods used in face recognition systems.

In Chapter 4, we introduced Densely Sampled Asymmetric Features (DSAF), which are capable of extracting sufficient discriminatory information to recognize face images and extract other useful facial information such as demographic informative features. The introduced feature representation forms a basis for subsequent face recognition and demographic analysis.

In Chapter 5, we introduced facial asymmetry based matching score space (MSS) based approach to recognize age-separated face images, which offered following contributions:

- Presented the first large scale facial asymmetry measurement and evaluation across temporal variations on three publically available face aging databases, the

FERET, MORPH, and FG-NET. The results revealed that facial asymmetry is an intrinsic facial characteristic which increases with age.

- Matching scores of DSAF features are combined with existing PCA and LBP based matching scores in a score space with SVM as classifier to recognize age-separated face images.
- The experimental results on FERET, MORPH, and FG-NET databases suggest that the proposed approach is more adaptable to recognize age-separated face images, compared to some existing state-of-the-art methods.
- Finally, the recognition results suggest that it is more difficult to recognize face images with large temporal variations compared to those with small temporal variations.

The study on facial asymmetry based age group estimation and its role in recognizing age separated face images presented in Chapter 6, offered the following contributions:

- Performed facial asymmetry based age group estimation using the benchmarks MORPH and FERET aging databases.
- Compared the performance of proposed age group estimation approach with state-of-the art method Face++. The results show superior performance of proposed approach in estimating age groups of query face images.
- Demonstrated the role of different asymmetric facial regions in recognizing age-separated face images.
- Integrated the knowledge learned from age group estimation of query face image into face recognition algorithm to enhance the recognition accuracy.
- Face recognition accuracies on MORPH and FERET databases suggest the superior performance of proposed face recognition approach compared to some existing state-of-the-art methods.

  In all face identification experiments the proposed approach outperforms the sum rule fusion applied on face and difference half face and three existing state-of-the-

art approaches, DM [59], MDL [62], and BFF [13], on both the MORPH and FERET datasets, for small and large temporal variations.

To analyze the errors introduced by age group estimation of probe image, we computed the rank-1 accuracy of face, difference half face, region 1, region 2, region 3 and region 4, for both scenarios: (i) with actual age groups and (ii) with estimated age groups. We observed a decrease in recognition accuracy of 3.32% and 2.65% for older and oldest image sets in case of MORPH datasets. In case of FERET, accuracy decreases by 3.44% and 3.67 % for dup I and dup II sets respectively. The minimal decrease in accuracy both for MORPH and FERET aging datasets show the superior age group estimation performance of proposed approach.

Finally, the study on facial asymmetry based demographic estimation and its role in recognizing age separated face images presented in Chapter 7 offered the following contributions:

- Performed AdaBoost feature selection of proposed DSAF features for demographic estimation of age group, gender, and race.
- Demonstrated the role of selected DSAF based features in estimating age group, gender, and race estimation. Comparison of results showed that proposed approach surpass the performance of state-of-the-art approach, Face++.
- Demonstrated the significance of different asymmetric facial regions in recognizing face images of different age groups, gender, and race of a query face image.
- Integrated the knowledge learned from demographic estimation of age group, gender and race into face recognition algorithm to enhance the recognition performance for age-separated face images.

Comparison of face recognition accuracies of three studies presented in this dissertation reveals that:

Firstly, the matching score space based approach (MSS) presented in Chapter 5 achieves face recognition accuracies of 70.08% and 63.24% on FERET, 75.40% and 69.40% on MORPH while 74.39% and 69.51% on FG-NET database for small and large temporal variations, respectively.

Secondly, the age group estimation base approach, introduced in Chapter 6 achieves the Rank-1 recognition accuracies of 78.10% and 73.01% for older and oldest image sets of for MORPH database, respectively. In case of FERET database, the Rank-1 recognition accuracies of 73.82% and 67.09% are achieved for dup I and dup II sets representing small and large temporal variations, respectively. The experimental results suggest that incorporating the knowledge learned form age group estimates improves the recognition performance compared to the matching score space based approach (MSS) introduced in Chapter 5.

Finally, the demographic estimation based approach introduced in Chapter 7, achieves the Rank-1 identification accuracies of 80.05% and 75% for older and oldest face image sets respectively on MORPH database. In case of FERET database, the proposed approach achieves a Rank-1 recognition accuracy of 77% for dup I representing small temporal variations. For the experiment, where images with large temporal variations (dup II set) are used as probe set, the proposed approach achieves a Rank-1 recognition accuracy of 75.21%. Recognition results suggest that incorporating the knowledge learned from demographic estimates including age group, gender and race further improves the recognition accuracy compared to both, the matching score space based approach (MSS) and age group estimation based approach presented in Chapter 5 and Chapter 6, respectively.

8.2 Future Directions

Though contributions of this dissertations were offered across a range of challenges in facial asymmetry based demographic estimation and recognition of age-separated

face images, yet there are new challenges both in recognizing age-separated face images and beyond. The following research directions appear promising.

In Chapter 5 a facial asymmetry based score space approach was presented to recognize age-separated face images. However the following research directions appear promising.

- Use of 3D facial aging databases to evaluate facial asymmetry and recognize age-separated face images
- We used holistic and local features along with asymmetric facial features to calculate matching scores to recognize age-separated face images. However, additional features may be considered for recognition task.
- Facial asymmetry may be introduced while modeling age progression in craniofacial growth models that characterizes growth related shape variations observed in human faces. Ensemble of classifiers such as a two-stage SVM classifier or an SVM based classifier followed by a Bayesian classifier in score space may be applied to enhance the recognition performance

In Chapter 6, our purpose was to integrate the knowledge learned from age group estimation into face recognition algorithm to enhance the recognition performance. The researchers may also take into account the following important aspects:

- Research on integrating the knowledge learned from facial age modeling into face recognition algorithm would be very beneficial to enhance the recognition performance across time lapse.
- Additional features such as morphological or color based features to encode facial marks, moles, and scars should be considered to recognize age separated face images. Since facial moles increases with increasing age, moles can be used as a strong indicator of age in applications such as age group estimation.

In Chapter 7, we integrated facial asymmetry based demographic estimates into face recognition algorithm to enhance recognition performance for age-separated face images. The study can be extended to following research directions as well.

- Human perception based fusion of age group, gender and race estimates into existing face recognition algorithms would be another research direction to be sought.
- The proposed demographic estimates including age group, gender and race can be combined to build a unified system for video based face recognition.

Perhaps the research on recognition of age-separated face images presented in this dissertation offers interesting and ambitious avenues for future research. Combination of different approaches to this problem should in turn yield a number of novel approaches that can be extended to standard face recognition approaches. One of which is application of asymmetric facial features to the problems outside the scope of face recognition. For example image retrieval scenarios such as matching database of large images with thumbnail images

REFERENCES

[1] S Z Li and A K Jain, *Handbook of Face Recognition*, 2nd ed. New York: Springer, 2011.

[2] M Moscovitch, G Winocur and M Behrmann , "Facing The Issue: New research shows that brain processes the faces and objects in separate brain systems," *Journal of Cognitive Neuroscience*, 1997.

[3] X Jiang et al., "Evaluation of shape-based model of human face discrimination using fMRI and behavioural Techniques," *Neuron*, vol. 50, 159-172 2006.

[4] P J Grother and M Ngan. (2014) NIST Interagency Report (8009), Available: http://biometrics.nist.gov/cs_links/face/frvt/frvt2013/NIST_8009.pdf, Acessed Sep 16, 2014

[5] P Phillips et al., "An introduction to the good, the bad, & the ugly face recognition challenge problem. ," *In Proceedings of Int. Conference on Automatic Face and Gesture Recognition*, 2011.

[6] A K Jain, B F Klare and U Park , "Face recognition: Some challenges in Forensics," *In Proceedings of Int. Conference on Automatic Face and Gesture Recognition*, 2011.

[7] M H Yang, D J Kriegman and N Ahuja, "Detecting faces in images: A survey," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 1, pp. 34-58, 2002.

[8] A K Jain, B Klare and U Park, "Face matching and retrieval: Applications in forensics," *IEEE Multimedia*, vol. 19, no. 1, pp. 20-28, 2012.

[9] M Turk and A Pentland, "Eigenfaces for recognition," *Journal of Cognitive*

Neuroscience, vol. 3, no. 1, pp. 71-86, 1991.

[10] P Belhumeur, J Hespanha and D Kriegman, "Eigenfaces vs. fisherfaces: Recognition using class specific linear projection," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 19, pp. 711-720, 1997.

[11] T Ahonen, A Hadid and M Pietikainen, "Face description with local binary patterns: Application to face recognition," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 28, no. 12, pp. 2037-2041, 2006.

[12] G Lowe, "Distinctive image features from scale-invariant kepoints," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91-110, 2004.

[13] D Yadav, M Vatsa R Singh, and M Tistarelli, "Bacteria foraging fusion for face recognition across age progression," in *IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pp. 173-179, 2013.

[14] D Yadav, R Singh, M Vatsa and A Noore, "Recognizing Age-Separated Face Images: Humans and Machines.," *PLoS ONE, doi: 10.1371/journal.pone.0112234*, vol. 9, no. 12, 2014.

[15] B Klare and A K Jain, "On a taxonomy of facial features," *Proc. IEEE Conf. Biometriics:* Theory, *applications and systems*, 2010.

[16] Skincare Physicians, Available: http://www.skincarephysicians.com/agingskinnet/basicfacts.html, Accessd: Sep 15, 2014.

[17] I Ercan et al., "Facial Asymmetry in young healthy subjects evaluated by statistical shape analysis," *Journal of Anatomy*, vol. 213, pp. 663-669, 2008.

[18] M Sajid, I A Taj, U I Bajwa and N I Ratyal, "The Role of Facial Asymmetry in Recognizing Age-Separated Face Images," *Computers and Electrical Engineering*,

(2016) http://dx.doi.org/10.1016/j.compeleceng.2016.01.001.

[19] M Sajid, I A Taj, U I Bajwa and N I Ratyal, "Facial Asymmetry based Age-Group Estimation: Role towards Recognizing Age-separated Face Images ," *Neural Computing and Applications (under Review)*, 2015.

[20] M Sajid, I A Taj, U I Bajwa and N I Ratyal, "Role of Facial Asymmetry based Demographic Estimation in Recognizing Age-separated Face Images," *IEEE Transactions on Information Forensis ans Security, (under review)*, 2016.

[21] P Viola and M J Jones. , "Robust real-time face detection," *International Journal of Computer Vision*, vol. 57, pp. 137-154, 2004.

[22] J Parris et al., "Face and eye detection in dataset," in *Intternational Joint Conference on Biometrics*, 2011.

[23] T F Cootes, G J Edwards and C J Taylor, "Active appearance models," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 6, pp. 681-685, 2001.

[24] T Cootes, C Taylor, D Cooper and J Graham, "Active shape models-their training and application," *Computer Vision and Image Understanding*, vol. 61, no. 1, pp. 38-59, 1995.

[25] Megvii I, "Face++ research toolkit," Available: www.faceplusplus.com, 2013.

[26] X Tan and B Triggs, "Enhanced local texture feature sets for face recognition under difficult lighting conditions ," *IEEE Transactions on Image Processing*, vol. 19, no. 6, pp. 1635-1650, 2010.

[27] K Ricanek Jr and T Tesafaye, "Morph: A longitudinal Image Database of Normal Adult Age-Progression," in *Int'l Conf. Automatic Face and Gestutre Recognition*,

pp. 341-345, 2006.

[28] H Hu and A K Jain, "Age, Gender and Race Estimation from Unconstrained Face Images," *MSU Technical Report Available: www.cse.msu.edu*, 2014.

[29] R Duda, P Hart and D Stork, *Pattern Classification*, 2nd ed.: Wiley-Interscience, 2000.

[30] B Heisele, P Ho and T Poggio, "Face recognition with support vector machines:global versus component-based approach," in *Proceedings of International Conference on Computer Vision*, 2001.

[31] A Ross and A K Jain, "Information Fusion in Biometrics," *Pattern recognition Letters*, vol. 24, no. 13, pp. 2115-2125, 2003.

[32] W Zhao , R Chellappa and P J Phillips and A Rosenfeld, "Facing Recognition: Literature Survey," *ACM Computing Surveys*, vol. 35, pp. 399-458, 2003.

[33] A F Andrea, M Nappi, D Riccio and G Sabatino, "2D and 3D face recognition: A survey," *Pattern Recognition Letters* , vol. 28, pp. 1885-1906, 2007.

[34] G Baudat and F E Anouar, "Generalized discrimnant analysis usin a kernel approach," *Neural Comutation*, vol. 12, pp. 2385-2404, 2000.

[35] L Wiskott, J Fellous, N N Kruger and C Malsbur, "Face recognition by Elastic Bunch Graph Maching," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, pp. 775-779, 1997.

[36] A Lanitis, C J Taylor, and T F Cootes, "Automatic interpretation and coding of face images using flexible models," *IEEE transactions on pattern analysis and machine intelligence*, vol. 19, no. 7, pp. 743-756, 1997.

[37] V Blanz and T Vetter, "A morphable model for synthesis of 3D faces," *In Proceedings of ACM SIGGRAPH*, pp. 187-194, 1997.

[38] N Ramanathan and R Chellappa, "Computational Methods for modeling facial aging: A survey," *Journal of Visual Languages and Computing*, vol. 20, no. 3, pp. 131-144, 2009.

[39] N Ramanathan and R Chellappa, "Modeling Age Progression in Young Faces," *Proceedings of IEEE CVPR*, pp. 387-394, 2006.

[40] J Suo et al., "Design Sparse Features for Age Estimation Using Hierarchical Face Model," *Proceedings of FGR*, pp. 1-6, 2008.

[41] Y Fu and T S Huang, "Human Age Estimation with Regression on Discriminative Aging Manifold," *IEEE Trans. Multimedia*, vol. 10, no. 4, pp. 578-584, 2008.

[42] K Luu, K Seshadri, M Savvides, T Bu and C Suen, "Contourlet Appearance Model for Facial Age Estimation," *Proceedings of IJCB*, pp. 1-8, 2011.

[43] J Bekios-Calfa, J M Buenaposada and L Baumela, "Revisiting Linear Discriminant Techniques in Gender Recognition," *IEEE Trans. Pattern Anal. Mach. Intell*, vol. 33, no. 4, pp. 858-864, 2011.

[44] T Wu, P Turaga, and R Chellappa, "Age Estimation and Face Verification across Aging Using Landmarks," *IEEE Trans. Inf. Forensics Security*, vol. 7, no. 6, pp. 1780-1788, 2012.

[45] G Guo and G Mu, "Joint Estimation of Age, Gender and Ethnicity: CCA vs. PLS," *Proc of FGR*, pp. 1-6, 2013.

[46] J E Tapia and C A Perez, "Gender Classification Based on Fusion of Different Spatial Scale Features Selected by Mutual Information from Histogram of LBP, Intensity, and Shape," *IEEE Trans. on Information and Forensics Security*, vol. 8,

no. 3, pp. 488-499, 2013.

[47] S E Choi, Y J Lee, S J Lee, K R Park and J K Kim, "Age Estimation Using a Hierarchical Classifier Based on Global and Local Facial Features," *Pattern Recogn*, vol. 44, no. 6, pp. 1262-1281, 2011.

[48] X Geng, C Yin and Z -H Zhou, "Facial Age Estimation by Learning from Label Distributions," *IEEE Transactions on Pattern Anal.Mach. Intell*, vol. 35, no. 10, pp. 2401-2412, 2010.

[49] H Han, C Otto and A K Jain, "Age Estimation from Face Images: Human vs. Machine Performance," *Proceedings of ICB*, pp. 1-8, 2013.

[50] H Han, C Otto, X Liu and A K Jain, "Demographic Estimation from Face Images: Human vs. Machine Performance," *IEEE Trans. Pattern Anal. Mach. Intell*, vol. 37, no. 6, pp. 1148-1161, 2014.

[51] Y Ylioinas, A Hadid and M Pietikainen, "Age classification in unconstrained conditions using LBP variants," *Proc. ICPR*, 2012.

[52] J Lu and Y Tan, "Ordinary Preserving Manifold for Human Age and Head Pose Estimation," *IEEE Transactions on Human-Mach. System*, vol. 43, no. 2, pp. 249-258, 2013.

[53] J Wang, Y Shang, G Su and X Lin, "Age simulation for face recognition," *Intl. Conf. Pattern recog.*, pp. 913-916, 2006.

[54] N Ramanathan and R Chellappa, "Face verification across age progression," *IEEE Trans. on image processing*, vol. 15, no. 11, pp. 3349-3361, 2006.

[55] X Geng, Z Zhou and K Smith-Miles, "Automatic age estimation on facial aging patterns," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 29, no.

12, pp. 2234-2240, 2007.

[56] S Biswas, G Aggarwal, N Ramanathan and R Chellappa, "A non-generative approach for recognition across aging," *IEEE International Conference on Biometrics Theory, Applications and Systems*, pp. 1-6., 2008.

[57] T Wu, P Turaga and R Chellappa, "Age Estimation and Face Verification across Aging Using Landmarks," *IEEE Trans.Inf.Forensic Security* , vol. 7, no. 6, pp. 1780-1788, 2012.

[58] U Park, Y Tong and A K Jain, "Age-invariant face recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 5, pp. 947-954, 2010.

[59] Z Li, U Park and A K Jain, "A discriminative model for age invariant face recognition," *IEEE Transactions on Information Forensics and Security*, vol. 6, no. 3, pp. 1028-1037, 2011.

[60] H Ling, S Soatto, N Ramanathan and D Jacobs, "A study of Face Recognition as People Age," *in Intl. Conference on Computer Vision*, 2007.

[61] H Ling, S Soatto, N Ramanathan and D Jacobs, "Face verification across age prgression using discrimnative methods," *IEEE Trans. Information Forensic and Security*, vol. 5, no. 1, pp. 82-91, 2010.

[62] D Sungatullina, J Lu, G Wang and P Moulin, "Multiview Discriminative Learning for Age-Invariant Face Recognition," in *10th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG)*, Shanghai , pp. 1-6., 2013.

[63] Y Liu, K Schmidt, J Cohn and S Mitra, "Facial asymmetry quantification for expression-invariant human identification," *Computer Vision and Image Understanding*, vol. 91, no. 1/2, pp. 138-159, 2003.

[64] N F Troje and H H Buelthoff, "How is bilateral symmetry of human faces used for recognition of novel views?," *Vision Research*, vol. 38, no. 1, pp. 79-89, 1998.

[65] S Gutta and H Wechsler, "Face Recognition using Asymmetric Faces," in *Ist International Conference on Biometric Authentication*, Hong Kong, 2004.

[66] J Harguess and J K Aggarwal, "Is there a connection between face symmetry and face recognition?," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 2011, pp. 66-73.

[67] A K Singh and G C Nandi, "Face recognition using facial symmetry," in *Proceedings of the Second International Conference on Computational Science, Engineering and Information Technology* , 2012, pp. 550-554.

[68] R Thornhill and S W Gangestad, "facial attractiveness," *Trans. Cognitive Sciences*, vol. 3, no. 12, pp. 452-460, 1999.

[69] D Hopea et al., "Symmetry of the face in old age reflects childhood social status," *Economics & Human Biology*, vol. 11, no. 2, pp. 236-244, 2013.

[70] C K Richardson, D Bowers, R M Bauer, K M Heilman and C M Leonard, "Digitizing the moving face during dynamic displays of emotion," *Neuropsychologia*, vol. 38, no. 7, pp. 1028-1039, 2000.

[71] J Chen, C Yang, Y Deng, G Zhang and G Su, "Exploring Facial Asymmetry using Optical Flow," *IEEE Signal Processing Letters*, vol. 21, pp. 792-795, 2014.

[72] NIST, "FERET Database," 2004. Available: http://www.itl.nist.gov/iad/humanid/feret, Acessed Sep 15, 2014.

[73] M Lades et al., "Distortion invariant object recognition on the dynamic link architecture," *IEEE transactions on computers*, vol. 42, pp. 300-311, 1993.

[74] T Serre, L Wolf and T Poggio, "Object Recognition with Features Inspired by Visual Cortex," *in Proceedings of IEEE CVPR*, pp. 994-1000, 2005.

[75] U Park and A K Jain, "Face matching and retrieval using soft biometrics," *IEEE transactions on information forensics and security*, vol. 6, no. 3, pp. 1028-1036, 2011.

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

[77] S Raudys and A Jain, "Small sample size effects in statistical pattern recognition: Recommendations for practitioners," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 13, no. 3, pp. 252-264, 1991.

[78] X Wang and X Tang, "Random sampling for subspace face recognition," *International journal of Computer Vision*, vol. 70, no. 1, pp. 91-104, 2006.

[79] P Comon, "Independent component analysis-A new concept?," *Signal Processing*, vol. 36, pp. 287-314, 1994.

[80] X He, S Yan, y Hu, P Niyogi and H -J Zhang, "Face recognition using laplacianfaces," *EEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, pp. 328-340, 2005.

[81] Y Liu, R L Weaver, K Schmidt, N Serban, and J Cohn, (2001), Technical Report. Available: https://www.ri.cmu.edu/pub_files/pub2/liu_yanxi_2001_2/liu_yanxi_2001_2.pdf, Aacessed Sep 15, 2014.

[82] Y W Cheong and L J Lo, "Facial Asymmetry: Etiology, Evaluation and Management," *Chang Gung Med. J.*, vol. 34, pp. 341-351, 2011.

[83] M Pesaresi and J A Benidiktsson, "A new approach for the morphlogical segmentation of high resolution satellite imagery," *IEEE Transactions on Geoscience and Remote Sensing*, pp. 309-320, 2001.

[84] M K Hu, "Visual Pattern Verification by Moment Invariants," *IEEE Trans. Information Theory*, 1962.

[85] C. -C Chang and C. -J Lin, "LIBSVM: A library for support vector machines," *ACM Trans. Intell. Syst. Tech*, vol. 2, pp. 1-27, 2011.

[86] Z Yang and H Ai, "Demographic classification with local binary patterns," *Advances in Biometrics Springer Berlin Heidelberg*, vol. 4642, pp. 464-473, 2007.

[87] G W Snedecor and W G Cochran, *Statistical Methods*, 8th ed.: Iowa State University Press, 1989.

[88] H Hotelling, "The most predictable criterion," *The journal of education pschology*, vol. 26, pp. 139-142, 1936.

[89] H Hotelling, "Relations between two variates," *Biometrike*, pp. 321-377, 1936.

[90] V N Vapnik, *Statistical Learning Theory*. New York: John Wiley, 1998.

[91] P Thukral, K Mitra and R Chellappa, "A Hierarchical Approach for Human Age Estimation ," *Proc. IEEE ICASSP*, pp. 1529-1532, 2012.

[92] T Hastie, R. Tibshirani and J. Friedman, *The Elments of Statistical Learning* Stanford, Springer, 2009.

[93] D G Altman and J M Bland, "Diagnostic tests. 1: Sensitivity and specificity," *BMJ 308*, vol. 1552, 1994.

[94] J Kittler, "Combining Classifiers: A theoratical framework," *Pattern Analysis and*

Applications, vol. 1, pp. 18-27, 1998.

[95] A E Mayes et al., "Environmental and lifestyle factors associated with percieved age in Chinese women," *PloS ONE,* vol. 5, no. 12: e15270, 2010.

[96] C S Morrison and B Z Phillips, "The Relationship between Age and Facial Asymmetry," American Association of Plastic Surgeons, 91st Annual Meeting, 2012.

[97] H Ha, S Shan, X Chen and W Gao, "A comparative study on illumination preprocessing in face recognition," *Pattern Recognition*, vol. 46, no. 6, pp. 1691–1699, 2013.

[98] Y Freund and R E Schapire, "A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting," *Journal of Comuter and System Sciences*, vol. 55, no. 1, pp. 119-139, 1997.