Pg Hj Mohd Yassin, DK Hayati Bte (2016) ***Novel Template Ageing Techniques to Minimise the Effect of Ageing in Biometric Systems.*** **Doctor of Philosophy (PhD) thesis, University of Kent.**

Novel Template Ageing Techniques to Minimise the Effect of Ageing in Biometric Systems

A Thesis Submitted to The University of Kent
For The Degree of Doctor of Philosophy
In Electronic Engineering

DK Hayati Bte Pg Hj Mohd Yassin

April 2016

Abstract

This thesis is concerned with the effect of ageing on biometric systems and particularly its impact on face recognition systems. Being biological tissue in nature, facial biometric trait undergoes ageing. As a result developing biometric applications for long-term use becomes a particularly challenging task. Despite the rising attention on facial ageing, longitudinal study of face recognition remains an understudied problem in comparison to facial variations due to pose, illumination and expression changes. Regardless of any adopted representation, biometric patterns are always affected by the change in the face appearance due to ageing. In order to overcome this problem either evaluation of the changes in facial appearance over time or template-age transformation-based techniques are recommended.

By using a database comprising images acquired over a 5-years period, this thesis explores techniques for recognizing face images for identify verification. A detailed investigation analyses the challenges due to ageing with respect to the performance of biometric systems. This study provides a comprehensive analysis looking at both lateral age as well as longitudinal ageing.

This thesis also proposes novel approaches for template ageing to compensate the ageing effects for verification purposes. The approach will explore both linear and non-linear transformation mapping methods. Furthermore, the compound effect of ageing with other variate (such as gender, age group) are systematically analysed. With the implementation of the novel approach, it can be seen that the GAR (Genuine Accept Rate) improved significantly when compared to typical face recognition system.

Acknowledgements

I am greatly indebted to express my profound sense of gratitude to my supervisors Dr. Sanaul Hoque and Dr. Farzin Deravi for their continuous encouragement, keen interest and constant guidance throughout the course of my research work. My conversations with them have been a source of great encouragement, inspiration and learning. Without the financial support from the Ministry of Education of Negara Brunei Darussalam, this research will not have been achievable.

I would like to express my gratitude to my parents. Especially to my mother, Imm Suhana, who has been a source of encouragement and inspiration to me throughout my life and taking care of my son when I was abroad. To my father, Pg Hj Mohd Yassin, who has always been there for me and I am thankful for everything he has done. It is my privilege to thank my husband, Hj Abd-Rahman, for showering me with endless practical and emotional support throughout my research period. I would also like to thank my sister, Norleha, for her endless love and support over the years and finally my late grandmother, who was wise, far sighted and owned the courage to open the door for my education regardless of her lack of literacy.

I am thankful to Dr. Asad Ali, for the continuous support and has been part of my source of motivation. Needless to say, I am also honoured to be able to learn from you. I would like thank all my friends, colleagues, EDA IT support and also my family members who have extended their moral support to accomplish this research.

Contents

List of Figures

List of Tables

Abbreviations

AMM	**A**ctive **A**ppearance **M**odel
BPNN	**B**ack **P**ropogation **N**etwork
CFBPNN	**C**ascade **F**orward **B**ack **P**ropogation **N**etwork
DET	**D**etection **E**rror **T**radeoff
EER	**E**qual **E**rror **R**ate
FAR	**F**alse **A**ccept **R**ate
FR	**F**ace **R**ecognition
FRR	**F**alse **R**eject **R**ate
GAR	**G**enuine **A**ccept **R**ate
KFA	**K**ernel **F**isher **A**nalysis
KPCA	**K**ernel **P**rincipal **C**omponent **A**nalysis
LBP	**L**inear **B**inary **P**attern
LDA	**L**inear **D**iscriminant **A**nalysis
LPQ	**L**ocal **P**hase **Q**uantization
LTP	**L**ocal **T**enary **P**attern
PCA	**P**rincipal **C**omponent **A**nalysis
ROC	**R**eceiver **O**perating **C**harateristic

CHAPTER 1

Introduction

This thesis is concerned with facial ageing for biometric systems. Face ageing in humans is the result of multi-dimensional process of physical, physiological, and social change, which affects considerably the appearance of a human face. Ageing-related appearance variation due to bone growth normally occurs throughout childhood and puberty, whereas skin-related effects principally appear in older subjects. Looking from the face recognition point of view, ageing of the face images of the same person brings confusion which degrades the system performance dramatically. In many practical systems (e.g., passport control, etc.), the time intervals between two acquired images can lead up to several years. This thesis suggests a novel template ageing scheme which will improve the overall results in face recognition systems.

A facial biometric verification system consists of two main processes; enrollment and verification. In enrollment, individual's face samples are captured and pre-processed, and the features are then extracted. These extracted features are labelled with user's ID called the "template", representing his identity. Verification mode verifies claimed identity by comparing input sample(s) to the enrolled template(s). The efficiency of the facial biometric system depends on the representative capability of the enrolled templates. Performance of these systems can be measured in terms of FAR and FRR.

1.1 Biometric systems

As the biometrics field has matured, practical system deployment has expanded, an understanding of the ageing process and managing ageing related impacts is becoming an increasingly important issue in terms of ensuring reliability of authentication processes based on biometric systems. The demanding needs of the security industry (e.g. the introduction of electronic passports) increased the interest in automatic biometric authentication and verification. Factors such as ethnicity, lifestyle, an individuals biological sex, living environment and physiologic changes that occur with time are some of the uncontrolled factors in human ageing as indicated by Biswas et al. [12]. However, an understanding of the ageing related effects in biometric system is a challenging task. This section will list some issues which affect and limit ageing related in biometric systems. Ageing is a continuous process and this process varies from one individual to another (i.e. different people age in different way).

Geng et al. [13] suggested that the age progression of an individual depends on their genes and can also be affected by external factors related to environment, gender, lifestyle, usage of cosmetics and so on. Another issue is that there is a lack of biometric databases suitable for ageing related analysis. Similarly, it is still relatively difficult to find a biometric database which contains the same age labels of subjects during the the same enrollment period and/or biometric samples with a regular and increasing time-lapse. These effects and limitations have a significant influence in the investigation of issues in biometric systems such as enrollment/re-enrollment strategies, processes for optimising user-system interactions, monitoring of usability factors, and understanding and managing ageing effects in biometric systems. Nevertheless, the impact of age progression on biometric systems has become a very important topic in the biometric field, and is currently generating a large volume of research.

The face recognition has been a very popular area of research. A few evaluation methodologies have been used to evaluate these algorithms as well as varieties of algorithms for face recognition. Researchers focus their attention on selecting discriminatory features that are least affected by within-person types of variations commonly encountered in different types of biometrics. Automated Face Recognition(FR) has become an important function in a wide range of security and surveillance application, involving computer networks, smart phones, tablets, IP camera, etc. For example, in the context of controlled access to critical information on computer networks systems, the face modality may allow for a continuous and non-intrusive authentication [14].

Given that biometric authentication systems are intended to be used for long time intervals, one of the most important aspects of their design is the permanence of biometric features to retain their characteristics over a period of time. In contrast, the use of features with low permanence will result in the development of systems that are likely to fail to operate accurately as the time interval between the initial enrolment of a user and the actual identification process increases. On the other hand, it is essential to use age-sensitive feature clues related to the age of a person for application related with estimating the age of a user. Therefore, it is essential to have the ability to assess the permanence of biometrics features so that appropriate features are selected for the right application.

Ricanek et al. [15] run face recognition experiment using the MORPH database. According to the results, correct recognition rates decreases as the age span between train and test images increases. In the work of Ricanek [16], a craniofacial growth model was proposed based on periorbital/periocular recognition. The data was collected using MORPH-II database and age groups were allocated using various classifications to produce EER performance.

According to Lanitis et al. [17], when dealing with increased age difference between faces in the train and test sets, a reduction of 10 percent was recorded in the performance of two face recognition algorithms . Similarly, both pose and illumination expression observed deterioration in the recognition. Lanitis et al. [18] also looked into developing a method for quantifying ageing effects on biometric templates, allowing in such a way the selection of appropriate features are use in specific application.

The published literature on adult head and face ageing is diverse, spanning topics related to anthropology and biology, computer and forensic science, and medicine and pathology. It is imperative to understand that patterns, rates and characteristics of ageing may be different due to culture and lifestyle (environment and gender i.e male and female), and biology (internal and external appearances of gender differences, ancestry or genetics, trauma and disease) as well as idiosyncratic features, such as frequency and extent of facial expressions.

A previous work focused on the literature on the ageing skull and face by Albert et al. [1], suggested that extrinsic ageing is considered as one of the main cause skin ageing exposure. It can be due to lifestyle such as drug use, smoking and/or diet adult age-related craniofacial changes comprise hard tissue and soft tissue changes. The soft tissue of the face age will begin as early as 20s. However, between the 40s and 50s the changes are most noticeable and continue to become more pronounced between the 50s and 60s. Some typical ageing related skin changes are shown in Figure 1.1. The loss of elasticity that, in some cases, affects the appearance of biometric feature (i.e. face) and affects the process of generating biometrics features (i.e. voice and movements) or inhibits the data acquisition process (i.e. fingerprints and palm prints).

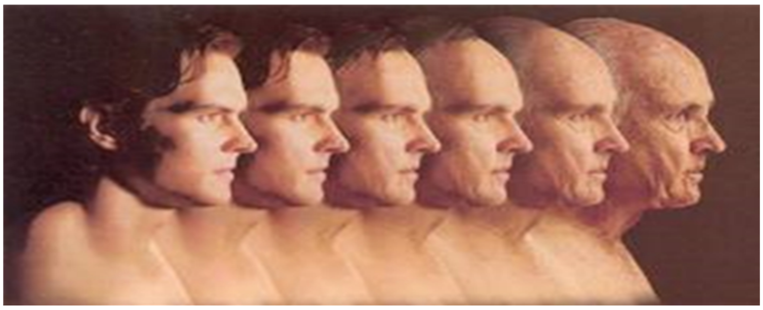

Figure 1.1 Typical ageing related in skin changes from early 20s to 60s [1]

1.2 Motivation

It is obvious that variation that affects all personal characteristics is caused by ageing. Face ageing is a challenging area and requires an in-depth understanding of the effects of individual changes towards a specific age span. People belonging to different ethnic groups may also have different age progression [19]. To date, the topic of generating biometric templates robust to ageing did not receive much attention. Face ageing variation causes modifications on biometric features that affect the matching between stored and captured biometric templates causing in that way deterioration in the performance of biometric authentication systems.

Figure 1.2 Challenges in face recognition: images of actress Jennifer Grey demonstrating variations due to pose, illumination, expression, and alterations through plastic surgery, ageing and makeup [2]

Conversely in the recent studies on face recognition, despite the rising attention to facial ageing, longitudinal study of face recognition remains an under-studied problem in comparison to facial variations due to pose, illumination and expression changes. shown in Figure 1.2. Most of the reported studies in relation to face ageing focused on age estimation [13], simulating the ageing process for each individual in face appearance as different age [20]. A commonly adopted solution in the state-of-the-art is the virtual template synthesis for ageing and de-ageing transformation involving complex 3D modelling techniques. However, these schemes are prone to estimation errors in the synthesis.

The ageing factor is very significant in face images. In comparison to other facial variation (pose, illumination, etc.), adaption to template ageing deserves a dedicated treatment of its own, since ageing is a life long process. Ageing also bring gradual changes in the data distribution over time, thus causing performance loss as a result of template becoming outdated. These factors indicate that template ageing process is very similar to the concept drift theory [21], based on the fact that real-worlds concepts change with the time resulting in underlying data distribution to change.

The challenging problem of facial ageing from a computer vision perspective can be described as follows:

- Diversity of Ageing Variation: Both the rate of ageing and the type of age-related effects differ for different individuals. Typically diverse ageing effects are encountered in subjects of different ethnic origins and different genders. External factors may also lead to diversities in the ageing pattern adopted by different individuals. Such factors include health conditions, lifestyle, psychological conditions, climatic related factors and deliberate attempts to intervene with the ageing process through the use of anti-ageing products [22] or cosmetic surgeries. As a result common ageing patterns cannot be applied successfully to all subjects.

- Intrinsic factors: Intrinsic ageing is caused by internal biological factors. Facial ageing is mainly due to the natural changes that occur as soft tissues lose their elasticity, muscle tone, and volume [23] as well as the facial bone shape modifications resulting from the lifelong and ongoing process of cranium remodeling [24]. Additional factors which affect these changes are an individuals biological sex, ethnicity, and features (i.e., features are purely unique to an individual such as facial expressions).

- Extrinsic factors: Extrinsic factors regulating facial ageing are due to lifestyle such as diet, drug use, and/or smoking [25] but the main cause of skin ageing is exposure to solar ultraviolet rays and the same is known as photoaging [26].

- Shape and texture: Generally, facial ageing causes both change in shape and texture of the human face. Since facial ageing introduces progressive variations in facial appearances, the characterizing models should account for the temporal nature of the induced variations. Therefore, by attributing facial shape and facial texture as a function of time, facial ageing may be modeled.

- Feature selection: To identify the appropriate form of data that provides a fair description in the FR process, important step of modelling facial ageing or age-invariant is required. The forms of data that can be assisted could be the template ageing where feature vectors are being extracted using specific feature selection. Also, the data could be individual specific or population specific.

- Controllability: Under normal mode the effects of ageing cannot be controlled and/or reversed unlike other variate such as pose, illumination, pose , expression and glasses etc. Thus, during face image capture, it is not possible to rely on the co-operation of the person for eliminating ageing variation. In addition, the process of collecting training data suitable for studying the effects of ageing requires long-time span intervals.

- Body composition: The muscle fibers and the fat deposition are generally reckoned as body composition changes with ageing in adulthood. With growing age, the muscle fiber atrophy and fat dissolve. It has been found that the appendages become thinner while abdominal girth expands due to general weakening of the abdominal muscles [27].

Figure 1.3 Face image of same individual, Coleen Nolan at age 44 [3]

Template ageing refers to the changes which may affect a specific system because the sample used at verification may differ from those obtained during enrollment (because of age-induced changes occurring during the time lapse). Figure 1.3 show a good visual example from the physical facial ageing perspective. The impact on the intra-class variations can often spoil the uniqueness of a face template even after a short time from the user's enrollment. It has been reported in several evaluation studies that even a short elapsed time of one or two weeks between acquisition sessions can induce variations in the face appearance which degrade the performance of the recognition system [28]. It is still not presently possible to compare these findings across other modalities on a like-for-like basis. So the main problem with the existing biometric system is that the performance of any face recognition algorithm drops just because the changes of a face impair the matching of the captured face image with a previously acquired reference face image, or template.

In order to overcome this problem, an investigation of the effects of ageing is thus crucial to optimal deployment and management of biometric systems. The algorithms proposed in this study are based on the age groups and age intervals as age progresses. The experiment requires the same subjects which can evaluate the changes in the facial appearance over time. For example, it is important to know if the effect of time lapse between enrolled and verified biometric data is similar in all age groups (i.e. at all physical ages) or if the recognition accuracy is seen to be degraded linearly or non-linearly as a function of the time difference under consideration. The main deficiency in analysis of these issues is the difficulty in finding the same individual subject which provides data over a long time lapse. Experiments on the face images of individuals over time show that the identification of a younger population is more difficult than for an older population in face recognition systems [15].

1.3 Aims and Objectives

The primary objective of the work to be presented in this thesis are to explore, improve the performance and consider how to manage the physical ageing in biometric systems. As noted in previous section template ageing is a relatively under-researched area which appear to show some inconsistencies in the results they present. The most effective methodology would be to follow a fixed population over an extended time period, rather than studying different population at different stages of ageing. The specific objective of this research is to investigate and improve the biometric systems based on different facial representations under the ageing effect. More specifically, extensive experimental studies have been carried out with the respect of other variate of ageing such as gender and age groups to reflect how the age of a sample may influence the verification error performance and how these can be manipulated in order to guide practical implementation towards achieving more reliable performance. In the latter experimental studies, template ageing using both linear and non-linear mapping methods are explored to further investigate the effectiveness of using feature transformation as a procedure for facial template ageing. The proposed method can approximate the process of ageing and hence can enhance the overall performance. This work will also explore the use of fusion techniques to improve the efficiency of the scheme and to optimise the proposed algorithms and carry out comparative analysis.

1.4 Structure of the thesis

This chapter has presented an initial overall review of previously reported research related to work about the ageing effects in biometric systems. It provides an important fundamental understanding of the types of ageing issues in biometric systems and current findings in terms of their impact in biometric system. The organization of the thesis is given below.

In Chapter 2, the background of various key concepts for understanding the related previous work on face ageing will be presented. A detailed comparative study is presented that focuses on various face recognition methods that had been reported in recent years. The related previous work has been divided into age progression and model-based face recognition.

Chapter 3 will provide details of the implementation and database that were used for training and testing purposes. This chapter will provide more detailed information of the subset used for the different types of ageing effect. The chapter will also cover the feature extraction approaches and feature projection techniques that are implemented in this research. This chapter will also explore and present quantitative results of a detailed investigation into the effects of ageing. It will discuss the effect of ageing in longitudinal and lateral face ageing.

Chapter 4 will discuss the performance evaluation of the implemented face recognition system to be used for the experimental study. This chapter will discuss and explore the different age groups and how each age groups perform better than the other. This chapter will further investigate other factors and how this information can be used in order to guide practical implementation towards achieving more reliable performance.

A novel template ageing proposed system using linear transformation will be presented in Chapter 5 and is aimed at improving the GAR performance as the time intervals increases. This chapter also explores other categories such as gender and age grouping using 2 types of face recognition system.

Chapter 6 will present another proposed template ageing system using non-linear transformation. This is to further enhance the accuracy of the proposed face recognition system using neural networks. This chapter will also further explore two different types of non-linear transformation models. A multi-classifier combining both algorithms with neural networks parameters will be investigate to produce more effective measures for face ageing system.

Chapter 7 will provide a final discussion of all the contributions made in this thesis. It will also provide list of recommendation of future works which may need to be carried out in order to further develop an improved strategies for managing ageing related issues in biometric system in the light of the finding in this thesis.

CHAPTER 2

Literature Review

2.1 Face recognition

A facial recognition system represents a computer-driven application that automatically authenticates a person from a digital image or a video sequence. It performs the recognition by comparing selected facial characteristics in the input image with a face database. Face recognition can be divided in following basic applications, although the used techniques are basically the same:

• Face identification: an unknown input face is to be recognized matching it against faces of different known individuals database. It is assumed that the person is in the database.

• Face Verification: an input face claims an identity and the system must confirm or rejects it. The person is also a member of the database.

• Watch List: an input face, presented to the system with no claim of identity, is matched against all individuals in database and ranked by similarity. The individual identity is detected if a given threshold is surpassed and if not is rejected.

For decades, face recognition has been a very popular modality for biometric applications. In the previous works, in order to improve the accuracy and computational complexity in a tightly controlled scenario, a variety of algorithms has been reported to address various real-world challenges in the process, such as, variations in pose [29], illumination [30], and so on. One of the most significant variation factor that contribute to face and degrade the face recognition system performance is the biological ageing. Some of the uncontrolled factors in human ageing [12] are ethnicity, lifestyle, gender, living environment, physiological changes that occur with time etc. In truth, human ageing is not an overnight phenomenon but as a rule is gradual occurring over a period of years.

The problem of automatic face recognition itself alone is a composite task that involves detection and location of faces in a cluttered background, facial feature extraction, and subject identification and verification [31]. Depending on the nature of the application, e.g., image acquisition conditions, size of database, clutter and variability of the background and foreground, noise, occlusion, and finally cost and speed requirements, some of the subtasks are more challenging than others.

The detection of a face or a group of faces in a single image or a sequence of images, which has applications in face recognition, is a challenging task and has been studied by many researchers [32]. Once the face image is extracted from the scene, its gray level and size are usually normalized before storing or testing. In some applications, such as identification of passport pictures or drivers' licenses, conditions of image acquisition are usually so controlled that some of the preprocessing stages may not be necessary.

One of the most important components of an automatic face recognition system is the extraction of facial features, in which attempts have been made to find the most appropriate representation of face images for identification purposes. The main challenge in feature extraction is to represent the input data in a low-dimensional feature space, in which points corresponding to different poses of the same subject are close to each other and far from points corresponding to instances of other subjects faces.

However, there is a lot of within-class variation that is due to differing facial expressions, head orientations, lighting conditions, etc., which makes the task more complex. Closely tied to the task of feature extraction is the intelligent and sensible definition of similarity between test and known patterns. The task of finding a relevant distance measure in the selected feature space, and thereby effectively utilizing the embedded information to identify human subjects accurately, is one of the main challenges in face identification. Typically, each face is represented by use of a set of gray-scale images or templates, a small-dimensional feature vector, or a graph. There are also various proposals for recognition schemes based on face profiles [33] and isodensity or depth maps [34].

2.1.1 Challenges

As biometrics field has matured, practical system development has expanded, and as many systems have now been in use for some time, an understanding of the ageing process and managing related impacts is becoming an increasingly important issue in terms of ensuring of authentication processes based on biometric systems.

Face recognition specific challenges:

- **Images taken years apart**: Facial appearance can considerably change over the years, driving to False Reject errors that can provoke denial of access to legitimate users.

- **Similar identities**: Biometric recognition systems are susceptible of False Accept errors where an impostor with a similar identity can be accepted as a legitimate user.

- **Biometric database**: There is a lack of biometric database suitable for ageing related analysis. For instance, it is still relatively difficult to find a biometric database which contains age labels of individuals (in addition to the identity labels) and/or biometric samples with regular and increasing time-lapse.

In some case, the training data used for face recognition systems are frontal view face images, but obviously, some difficulties appear when trying to analyze faces with different pose angles. Therefore, some processing schemes introduce rotated images in their training data set and other try to identify particular features of the face such as the eyes, mouth and nose, in order to improve positive ratios [11].

These effects and limitations have a significant influence in the investigation of issues in biometric systems such as enrolment strategies, understanding and managing ageing in biometric systems.

2.2 State-of-the-art

Traditionally age progression is performed by forensic artists [35] who make use of face ageing theory as described in anthropometric studies [36] in combination with images of relatives of a subject in order to create, either by hand or by using computer tools, age progressed drawings. In the case of automated age progression, the majority of age progression methods reported in the literature are data-driven.

2.2.1 Age progression

When dealing with increased age difference between faces in the train and test sets, Lanitis [17] observed that the performance of face recognition system drops by about 12%, on average, when dealing with test faces showing different types of variation other than the one in the training set such as ageing variation in upper face, nose, lower cheek and face only. Similarly, both pose and illumination expression caused deterioration in the recognition rate. The effects of ageing have greater impact when using face templates that include the upper face region compared to the lower part of the face and/or the nose. When using templates based on the lower part of the face ageing effects cause smaller decrease in the recognition performance. However, the regions affected more intensively by ageing are at the same time the most discriminatory regions.

Lantitis et al. [37] also reported results of an independent age progression. The experiments involved 3 tests where two of them were machine-based and another one was human-based. In test 1, the image similarity used methods such as AAM [38], LBP [39], Spatial Histogram of key-points (SHIK) [22], Histogram of oriented Gradients (MHOG) [40]. In test 2, classification test, where a reference dataset include 30 images from age between 10 to 75 years old. For test 3 which is the human based test, the study involved human experts to select the target images.

Lee et al. [41] proposed facial ageing and facial asymmetry based approaches to identify twins. Both approaches rely on the Modified Active Shape Model (MASM) approach [42] to automatically localize a dense set of facial landmarks around which feature are extracted using the Self-Quotient (SQ) algorithm [43]. Both approaches were thoroughly evaluated under 4 test scenario across 3 time-lapse of images (same day in 2009) with 96.9% accuracy rate, second collection (same day in 2010) with accuracy rate of 96.6% and data collections from 2009-2010 with 85.4%.

Lanitis et al. [44] proposed a framework using learned age transformation that can be used for simulating ageing effects on new face images in order to predict how individual might look like in the future. The three approaches that were used to build a statistical face model in the ageing classifier are (i) quadratic function of the model parameters based on the distribution, (ii) supervised and unsupervised neural networks trained onto the model parameter and (iii) characterised facial ageing effects using regression function. In their experiments, they used a private database containing 500 face images of 60 subjects. The results show that, there is evidence at 10% significance level to support the improvement in the performance when the weighted appearance-specific methods were used and evidence at 5% significant level when weighted person-specific method was used.

Burt et al. [45] created composite faces for different age groups by computing the average shape and texture of human faces that belong to these age groups. By incorporating the differences between composite faces on regular faces, they observed a change in the perceived age of faces. Gandhi [46] designed a support vector machine based age estimation technique and extended the image based surface detail transfer approach to simulate ageing effects on faces.

Tiddeman et al. [47] focused on the shape of each face by a set with 179 points located along contours around the major facial features (eyes, nose, and mouth) along the facial border using Active Shape Model [48]. To improve the textures of the facial prototypes, they proposed a wavelet-based method (see Stollnitz et al. [49] for an introduction on using wavelets in computer graphics). In the experimental validation of age transformation, they transformed 10 female and 10 male faces evenly divided into two age categories (18 to 45 and 45 years old). The increase in age produced by transformation with texture processing (mean 30.0 years, standard error ±1.3 years) corresponds to that expected from the age difference between the younger and older adult prototype faces (mean 27.9, standard error ±1.4 years). Hence, the wavelet-enhanced age transformation projected the faces by the correct amount and reached the target age bracket.

Ricanek et al. [50] examined the impact of normal adult ageing of the craniofacial region on the performance of the standard PCA Face recognition algorithm originally proposed by Turk [51]. They demonstrated the impact of ageing on image based and appearance based. Face Recognition techniques through simple image processing using MORPH-I and FERET databases. In their discussion, Rank-1 of MORPH presented an odds ratio of 0.6707, which means that as they increased the time difference by one year, it is expected a 0.6707 times change in the recognition rate. This shows that as the time difference increases their recognition rate decreased.

Rowland et al. [52] propose an age progression method based on age prototypes generated by averaging faces belonging to the same age group. Differences between age prototypes corresponding to different ages can be achieved by adding on a given face image in age progression. In a more recent approach by Kemelmacher-Shlizerman et al. [53], derived 40000 images collected from the internet and discussed an age progression method based on age prototypes. A key element of the proposed method is the alignment and illumination normalization used that ensures the generation of accurate age prototypes despite illumination, pose and expression variation in the training images.

Sethuram et al. [54] defined a hierarchical model for idiosyncratic facial ageing (i.e. features purely unique to the individual such as hyperdynamic facial expressions) and developed a synthetic facial ageing system based upon hierarchical models. The results showed that synthetic images generated from the hierarchical models consistently have a better match score than those generated by the general model. From his research, the score for an average of 8 subjects increases from 0.5265 with general model to 0.7905 with hierarchical model.

Sethuram et al. [55] demonstrated the severity of performance loss due to adult ageing and discuss steps towards improving the robustness of face-based biometrics using baseline PCA face recognition algorithm from the Colarado State Universitys Face Identification Evaluation System [56]. The dataset, comprising of both the FG-NET and MORPH-I Album was used for training and testing the algorithms. While using the original images only have Rank-1 accuracy of 18.75%, whereas when the experiments includes Age-Progressed Images, the Rank-1 accuracy increased to 31.25%. The influence of each axis during the age progression process was determined by maximizing the probability that an age progressed face belongs both to the distribution of faces at the target age and the distribution of differences between age-progressed samples and the actual faces of the same subject at the target age.

Ramanathan at al. [57] proposed a Bayesian age-difference classifier built on a probabilistic eigenspaces framework to perform face verification across age progression. The experimental results illustrated the applications of their proposed facial ageing model using Principal Component Analysis [51]. The performance for the Rank-1 recognition score under the following three setting: • No transformation in shape and texture achieved accuracy of 38%

• Performing shape transformation achieved accuracy of 41% and performing shape and textural transformation achieved accuracy of 51%.

Existing age progression algorithms are capable of generating face images with ageing effects. Actually there are various web applications (Ageing Booth [58]) that are designed for adding ageing related transformations. However, the key issue in automated age progression is not the generation of aesthetically pleasing results but the generation of accurate predictions. Age progression algorithms are not yet to a stage to produce highly accurate predictions for different subjects mainly due to the diversity of ageing effects, the dependence of the ageing process on external factors/events that may occur in a subjects life and also due to the compounded effect of ageing variation with other within-subject variations (i.e. expression, pose and illumination).

In some cases age progression is regarded as a missing data recovery problem. For example Wang et al. [59] treat the problem of age progression in the framework of a super resolution methodology. During the age progression process the input face is down sampled to low resolution and a super-resolution algorithm trained using images at the target age is used for creating a high resolution image that incorporates ageing effects on top of the basic id structure exhibited in the down-sampled image. In [60] the problem of age progression is treated as an occlusion removal problem where the appearance of a given face in conjunction with the recursive PCA method is used for predicting the appearance of a given face at the target age.

The lack of standardized age progression performance evaluation metrics in combination with the fact that in most cases researchers used different training/test sets, prevents the presentation of a comprehensive set of comparable age progression results [61]. Table 2.1 present typical evaluation methods used by different researchers.

Table 2.1 A summary of age progression evaluation methods

Publication Yr	Reference	Age Progression Method	Evaluation Method
2006	Ramanathan [20]	Craniofacial growth model	Face Recognition
2006	Scandrett [62]	Ageing Trajectories	Shape difference,texture difference
2007	Geng [13]	Ageing pattern subscape in AAM space	Mahalanobis distance , Face recognition results
2008	Lanitis [4]	Comparison of three methods (prototypes, ageing functions and SVM-distance-based)	Distance from id and age target distributions
2009	Sethuram [55]	AAM-based analysis by synthesis	Face Recognition
2010	Park [63]	3D face ageing model	Face Recognition
2012	Suo [64]	Concatenation Graph Evolution Ageing Model	Age similarity(user evaluation and age estimation algorithm ID similarity (user evaluation and face recognition algorithm)
2010	Luu [65]	Spectral Regression	Face Recognition
2012	Wang [59]	Tensor Space Analysis and AAM	Age similarity(user evaluation and age estimation algorithm), ID similarity (user evaluation and face recognition algorithm)
2014	Tsai [66]	Guided Prediction	Shape difference, user evaluation
2014	Kemelmacher-Shlizerman [53]	Illumination invariant age prototypes	Human based (Using Mechnical Turk)

2.3 Review of ageing effects in biometric system

When compared with other source of variation in face images, ageing variations is specific to a given individual. It can occur slowly and is affected significantly by other factors. The appearance of a human face is affected considerably by the ageing process. Examples of ageing effects on faces are shown in Figure 2.1. Facial ageing is mainly attributed to bone movement and growth, and skin related deformations. Ageing related appearance variation due to bone growth usually takes place during childhood and puberty whereas skin related effects mainly appear in older subjects. Skin related effects are associated with the introduction of wrinkles caused by reduced skin elasticity and reduction of muscle strength [22, 35]. There are few factors that are affecting the face recognition:

Figure 2.1 Face images displaying ageing variation. Each row shows images of the same individual [4]

• Gender: Gender differences in health, socioeconomic status, and social resources persist into advanced old age and result in variations in life trajectories and responses to the challenges of longevity [67]. Demographic changes have forced gerontologists to focus attention on the gender based character of population ageing [68]. The ageing process of male and female faces share many common features, attention to the particular differences in the ageing man is obvious [69]. Gender differences in the male face include the presence of facial hair, increased facial vascularity, increased thickness, increased fat content, hormonal influences, and potentially differing rates of fat and bone absorption during the life cycle. Women tend to develop more and deeper wrinkles in the perioral region than men; their skin contains a significantly smaller number of appendages than men [70]. Women who look young for their age have large lips, avoid sun- exposure and possess genetic factors that protect against the development of gray hair and skin wrinkles [71]. High social status, low depression score and being married are associated with a younger look, but the strength of the associations varies between genders [72].

- Unconstrained Environment and/or Health: A number of factors contribute to and in many instances accelerate the natural ageing process. Prolonged or frequent exposure to environmental agents such as sunlight (ultraviolet radiation) and wind or arid climates can cause skin, particularly the more delicate skin of the face, to age prematurely [73]. In addition to the photo damage caused by sunlight, which dries and destroys the cells and underlying structure of the skin, exposure to the sun gives the skin a furrowed, thickened appearance and hastens the development of wrinkles, especially around the eyes as a product of squinting [74]. Aridity and wind likewise dehydrate the skin, contributing to the formation of wrinkles, but the effect of ultraviolet radiation from the sun on the facial tissues by far exceeds these agents in effect. Chronic sun exposure can result in numerous changes in human skin, particularly on the face, changes in photo-ageing include wrinkling, elastosis, actinic keratoses, irregular pigmentation [75]. With continued exposure to the sun and other elements, the color and texture of the face can change, becoming blotchy, yellowish, and leathery, with loose, inelastic, hyper pigmented skin. Blood vessels lying close to the surface of the skin may become prominent as networks called spider veins, adding to the skin's overall mottled and blotched appearance [76].

Findings from [77] concluded that increased skin wrinkling was observed in the order: Caucasian > Hispanic >African American > East Asian. Similarly, although the ageing process of the male and female face share many common features, discrete anatomic differences exist between the male and female face. In the literature found in plastic surgery, some suggest the existence of differences in ageing due to gender differences. This also has been discussed in ageing changes in the male face [69].

- Ageing: Adult ageing is a distinct process from childhood growth and development. While in the formative years, pre-adult process leads to shape-based changes in the mid and lower face, adult ageing consists of shape deformation from weight change, minor bone remodeling, and tissue degeneration. Significant textural changes due to a variety of complex phenomenon occur due to adult ageing [1].

Physical ageing is the term used to describe the general effects of ageing process across the whole ageing cycle (i.e. related to the absolute biological age of an individuals) on the biometric measurements used in the verification process, without special reference to the time lapse between enrolment and verification. Changes in facial appearance due to ageing can affect discriminatory facial features, resulting in deterioration of the ability of humans and machines to identify aged individuals shown in Figure 2.2 .

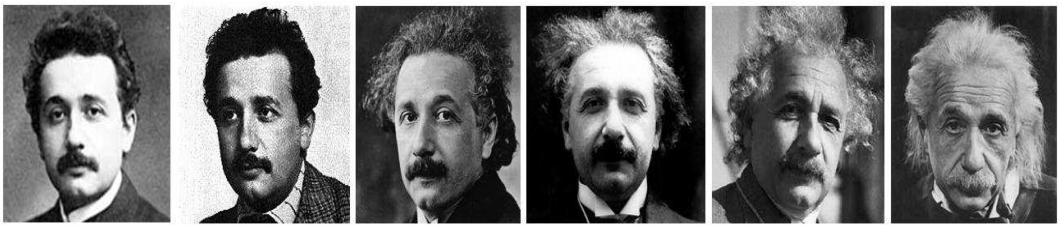

Figure 2.2 Albert Einsteins face ageing (collected by Internet image search [5]).

There are different ways in which ageing related changes appears such as change in shape of facial features (e.g., mouth, cheeks, or sagged eyes), weight loss and gain, wrinkles and speckles [78]. Due to these variations the performance of face recognition decreases. Though the template images can be updated, but there are different cases in which updating the template is not always possible such as screening, missing child problem and multiple enrollment problems. In these cases either they are trying to hide their identity or they might not be available. Due to all these conditions, the problem of ageing in face recognition is still an important area of research.

Another types of ageing is called the intrinsic ageing. Intrinsic ageing is caused by internal biological factors whereas extrinsic ageing is caused by environmental influences. Intrinsic factors causing facial ageing are mainly due to the natural changes that occur as soft tissues lose their elasticity, muscle tone, and volume [79] as well as the bony shape modifications resulting from the lifelong and ongoing process of bone remodeling [80]. Affecting these changes are an individuals biological sex, ethnicity, and idiosyncratic features (i.e., features purely unique to the individual such as hyperdynamic facial expressions).

On the other hand, extrinsic factor influencing facial ageing is lifestyle such as diet, drug use, and/or smoking (although this is debated [25]) but the main cause of skin ageing is exposure to solar ultraviolet rays known as photoageing [81].

Most of the previous researchers used different subjects on age intervals experiments and this eventually contributed to an unfair comparison towards all age groups. In this research, same subjects are being introduced and implemented for all age intervals. This method will address the above limitation that was done by previous researcher.

2.4 Face recognition techniques

Face recognition by a computer system is a very popular and useful application of (digital) image analysis. It is a method to identify or verify the identity of a person. In an identification process, based on face recognition, the face image of an unknown identity is compared with face images of known individuals from a large database. In the ideal case the system returns the determined (and of course the correct) identity. In a verification process the face image of a person is compared with face images from a database with the claimed identity. The system returns a value, which is a measure for the similarity between the two images. Of course this value needs to reach a certain (and in advance established) threshold to acknowledge or reject the claimed identity.

An example of an application of face recognition is person identification or verification at the customs authorities on an airport. A face recognition system could replace current identification methods like PIN-codes, passwords and ID-cards, but also extremely reliable methods of biometric person identification, like fingerprint analysis and retinal or iris scans. The disadvantage of methods like these is that they rely on the cooperation of the participants, whereas a person identification system based on the analysis of (frontal) images of the face can be effective without the participants cooperation or knowledge.

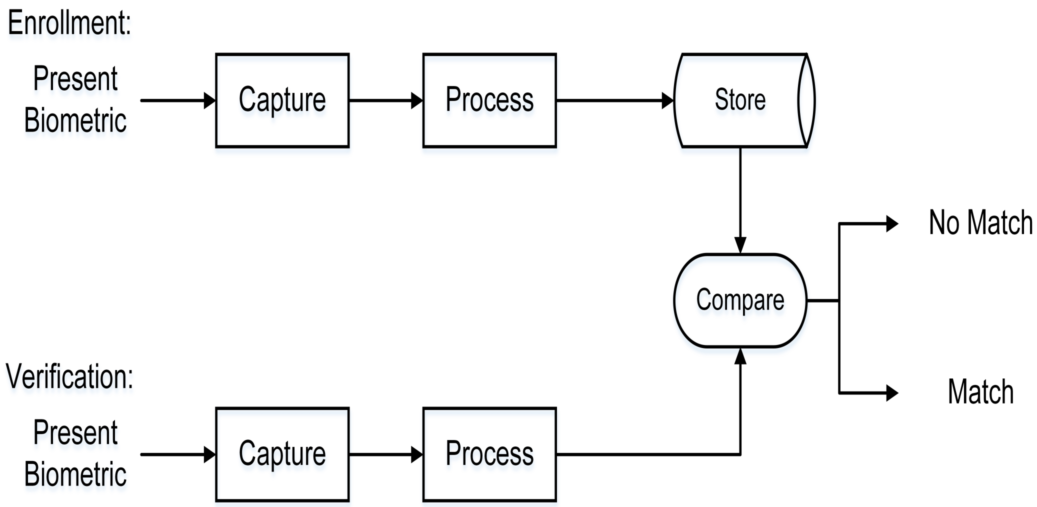

Figure 2.3 Principle of an verification process with face recognition

The process of person verification by using face recognition can be split into three main phases (Figure 2.3). These are registration and normalization, feature extraction and classification. In the registration and normalization phase, the image is transformed (scaled and rotated) till it has the same position as the images from the database. For example, this may mean that the eyes are at the same positions. In this part the problem factors like illumination differences are also reduced. In the feature extraction phase, the most useful and unique features (properties) of the face image are extracted.

With these obtained features, the face image is compared with the images from the database. This is done in the classification phase. The output of the classification part is the identity of a face image from the database with the highest matching score, thus with the smallest differences compared to the input face image. Also a threshold value can be used to determine if the differences are small enough. After all, it could be that a certain face is not in the database at all.

There are many feature extraction algorithms. Most of them are used in other areas than face recognition. Researchers in face recognition have used, modified and adapted many algorithms and methods to their purpose. Feature extraction methods in face recognition systems can be categorized into three types:

1. Generic methods - these methods are based on edges, lines and curves in an input image

2. Template based methods - these methods are used to detect the actual facial features of the face such as the eyes, nose, mouth

3. Structural matching methods - these methods take geometrical constraints on the features into consideration.

According to Jain et al. [82], there are three concepts that are key in building a classifier: similarity, probability and decision boundaries. We will present the classifiers from that point of view.

- Similarity: This approach is intuitive and simple. Patterns that are similar should belong to the same class. This approach have been used in the face recognition algorithms implemented later. The idea is to establish a metric that defines similarity and a representation of the same-class samples.

For example, the metric can be the Euclidean distance. The representation of a class can be the mean vector of all the patterns belonging to this class. The 1-NN decision rule can be used with this parameters. Its classification performance is usually good. This approach is similar to a k-means clustering algorithm in unsupervised learning. There are other techniques that can be used. For example, Vector Quantization, Learning Vector Quantization or Self-Organizing Maps - see Table 2.2. Other example of this approach is template matching. Researchers classify face recognition algorithm based on different criteria. Some publications defined Template Matching as a kind or category of face recognition algorithms [83]. However, we can see template matching just as another classification method, where unlabeled samples are compared to stored patterns.

Table 2.2 Similarity-based classifiers

Method	**Notes**
Template matching	Assign sample to most similar template. Templates must be normalized
1-NN	Eigenvector-based , non-linear map, uses kernel methods
(Learning) Vector Quantization methods	Assign pattern to nearest patterns class There are various learning methods.
k-NN	Like 1-NN, but assign to the majority of k nearest patterns.
Nearest Mean	Assign pattern to nearest class mean.
Self-Organizing Maps (SOM)	Assign pattern to nearest node, then update nodes pulling them closer to input pattern
Subspace Method	Assign pattern to nearest class subspace

- Probability: Some classifiers are build based on a probabilistic approach. Bayes decision rule is often used.

• Decision boundaries: This approach can become equivalent to a Bayesian classifier. It depends on the chosen metric. The main idea behind this approach is to minimize a criterion (a measurement of error) between the candidate pattern and the testing patterns. One example is the Fishers Linear Discriminant (often FLD and LDA are used interchangeably). Its closely related to PCA. FLD attempts to model the difference between the classes of data, and can be used to minimize the mean square error or the mean absolute error. Other algorithms use neural networks. Multilayer perceptron is one of them. They allow nonlinear decision boundaries. However, neural networks can be trained in many different ways, so they can lead to diverse classifiers. They can also provide a confidence in classification, which can give an approximation of the posterior probabilities. Assuming the use of an Euclidean distance criterion, the classifier could make use of the three classification concepts here explained.

A special type of classifier is the decision tree. It is trained by an iterative selection of individual features that are most salient at each node of the tree. During classification, just the needed features for classification are evaluated, so feature selection is implicitly built-in. The decision boundary is built iteratively. There are well known decision trees like the C4.5 or CART available. See Table **??** for some decision boundary-based methods, including the ones proposed in [82].

Other method widely used is the support vector classifier. It is a two-class classifier, although it has been expanded to be multi-class. The optimization criterion is the width of the margin between the classes, which is the distance between the hyperplane and the support vectors. These support vectors define the classification function. Support Vector Machines (SVM) are originally two-class classifiers. Thats why there must be a method that allows solving multi-class problems. There are two main strategies:

• On-vs-all approach: A SVM per class is trained. Each one separates a single class from the others [84].

• Pairwise approach: Each SVM separates two classes. A bottom-up decision tree can be used, each tree node representing a SVM [85]. The coming faces class will appear on top of the tree. Other problem is how to face non-linear decision boundaries. A solution is to map the samples to a high-dimensional feature space using a kernel function [86].

Different strategies and schemes have been attempted to solve the face recognition problem. The classification of these approaches into different categories is not easy because different criteria can be taken into account. But one of the most popular and genera classification schemes is the following one:

• Holistic methods: Based on overall information in the face. This kind of methods try to recognize faces in an image using the whole face region as an input, treating the face as a whole. For instance, statistical approaches use statistical parameters to create a specific face model or space to be used for the recognition stage. The training data is used to create a face space where the test faces can be mapped and classified.

• Feature-based methods: Non-holistic methods based in identifying structural face features such as the eyes, mouth and nose, and the relations between them to make the final decision. Early feature matching strategies were based on the manually definition of facial points and the computation of the geometric distances between them, resulting in a feature grid for people identification. Recent evolved algorithms place the facial points automatically and create an elastic graph (Lades et al. [87]) in order to resolve previous algorithms insensitivity to pose variations.

• Geometrical feature matching: The overall geometrical feature of a face is sufficient for recognizing an individual. The facial configuration is represented by the position and the size of the facial features like eyebrows, eye, nose, mouth and the face outline [83, 88–90]. Photometric approaches are statistical techniques that distill an image into values and compare these values with templates [91].

Mixture-distance based technique was introduced by Cox et al. [90], which achieved the accuracy of 95%. Thirty manually extracted distances represented every faces. B.S. Manjunath et al. [92] generated 35-45 feature points per face using Gabor Wavelet

Decomposition. After compensating the different centroid location, two cost values, the similarity cost and topological cost were evaluated. The recognition accuracy, which is the best match to the right individual, was 86% and 94% of the correct person faces.

The distance between features might be more useful for finding the matches in the database. But they are more dependent upon the accuracy of the feature dependent algorithm. Currently no algorithm based on face feature location gives a good performance and their computational time is more.

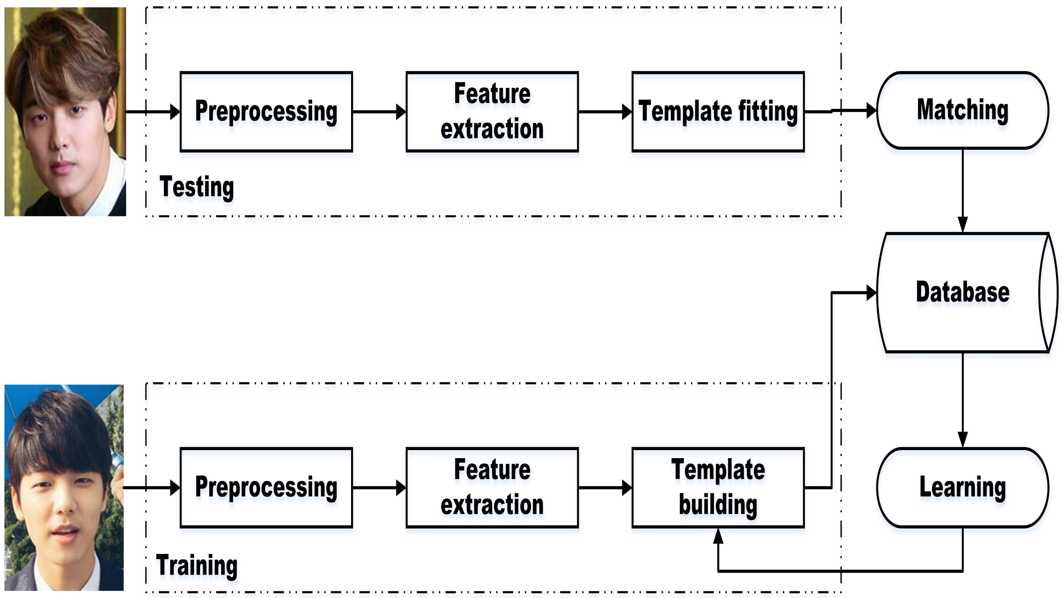

Figure 2.4 Template-matching algorithm diagram [6]

• Template matching methods: These algorithms compare input images with stored patterns of faces or features. Template matching methods try to define a face as a function. It is to find a standard template for all the faces. Different features can be defined independently. Patterns are represented by samples, models, pixels, curves and textures. The recognition function is usually a correlation or distance measure. Figure 2.4 is an example depicting a template-matching algorithm.

In template matching, the image is represented by the 2D arrays of intensity and the values are compared using metric such as Euclidean distance using a single template which represents the complete face. There are many techniques based on template matching, in which the complete face is represented by more than one template [93].

Bruneli et al. [83] selected four feature templates (eyes, mouth, nose and the whole face). While comparing the performance of the geometrical matching algorithm with the template matching algorithm on the same set of database, the team found that the geometrical matching algorithm is less superior as compared to template matching algorithm. Computational complexity is the biggest drawback of template matching and moreover the problem also lies in the template description. There are few discrepancies between the image which is to be tested and the template. This tolerance might affect the performance of the face recognition technique.

For example, a face can be divided into eyes, face contour, nose and mouth. Also a face model can be built by edges. But these methods are limited to faces that are frontal and unoccluded. A face can also be represented as a silhouette. Other templates use the relation between face regions in terms of brightness and darkness. These standard patterns are compared to the input images to detect faces. This approach is simple to implement, but its inadequate for face detection. It cannot achieves good results with variations in pose, scale and shape. However, deformable templates have been proposed to deal with these problems

- Appearance-based methods: A template matching method whose pattern database is learnt from a set of training images. The templates in appearance-based methods are learned from the examples in the images. In general, appearance-based methods rely on techniques from statistical analysis and machine learning to find the relevant characteristics of face images. Some appearance-based methods work in a probabilistic network. An image or feature vector is a random variable with some probability of belonging to a face or not. Another approach is to define a discriminant function between face and non-face classes. These methods are also used in feature extraction for face recognition as follows:

1. Eigenface-based: Sirovich et al. [94] developed a method for efficiently representing faces using PCA. The team's goal is to represent a face as a coordinate system. The vectors that make up this coordinate system were referred to as

eigen pictures. Later, Turk et al. [51] used this approach to develop a eigenface-based algorithm for recognition. PCA is a tool using which these eigenfaces can be extracted from an image. It is viceversa, that is, if a system is having a set eigenfaces then the original image of the face can be reconstructed [51, 94].

This approach is efficient and simple as compared to different techniques in a constrained environment. But this method has few limitations over unconstrained conditions such as expressions in which there is a minor change in the facial feature recognition and shape. Secondly, the difference in pose leads to the distortion of the distance of the objects. Thirdly the variations of the lighting conditions, where for example the bright light will cause an image saturation. But the lighting conditions can be overcome using the Fisherface method. It is an enhancement of eigenfaces but it uses the Fisher's LDA [95, 96].

For example, PCA was invented by Karl Pearson in 1901 [97], but was proposed for pattern recognition 64 years later [98]. Eventually, it was applied to face representation and recognition in the early 90s [51, 94, 99]. See Table **??** extraction algorithms for a list of some feature extraction algorithms used in face recognition.

The experiments with face recognition and ear, by using the standard PCA approach, showed that the recognition performance is identical using the ear images or the face images, but after combining the two for multimodal recognition, the result showed a significant improvement in the performance of 90.9% [100].

2. Distribution-based: These systems were first proposed for object and pattern detection by Sung [101]. The idea is to collect a sufficiently large number of sample views for the pattern class the team wish to detect, covering all possible sources of image variation.

3. Neural Networks: This technique (Viennet et al. [102]) can be considered as statistical approach (holistic method), because the training procedure scheme

usually searches for statistical structures in the input patterns. Some early researches used neural networks to learn the face and non-face patterns [103]. They defined the detection problem as a two-class problem. The real challenge is to represent the images not containing faces class. Other approach is to use neural networks to find a discriminant function to classify patterns using distance measures [101]. Some approaches have tried to find an optimal boundary between face and non-face pictures using a constrained generative model [104].

For the facial representation-based face recognition, there are several types of representation which are as follows:

• LBP [39] and Multi-scale Local Binary Patterns (MLBP) [105]: Ahonen et al. introduced a facial representation based on LBP texture descriptor. It is an efficient descriptor that assigns a label to every pixel of an image by thresholding the neighborhood of each pixel with the center pixel value and considering the result as a binary number. Then the histogram of the labels is obtained and used as a descriptor. Dissimilarity measure between the pair of facial images is obtained using χ^2 histogram distance. MLBP is an extended version of LBP by using multiple radii and offering the advantage of scale invariance.

• LPQ [106]: LPQ is based on quantizing the Fourier transform phase in local neighborhoods. The phase has been shown to have a blur invariant property under certain commonly fulfilled conditions. These descriptors are obtained for facial images in the manner similar to LPB and matched using χ^2 histogram distance [106].

• LTP [106]: These descriptors utilize the idea that many facial regions are relatively uniform, it is potentially useful to improve the robustness of the underlying descriptors in these areas.

• Elastic Bunch Graph Matching (EBGM) [107]: Elastic Bunch Graph Matching localizes a set of landmark features and extracts Gabor jets at landmark positions.

- Scale Invariant Feature Transform (SIFT) [108]: Scale Invariant Feature Transform (SIFT) are invariant to image scaling, translation, rotation, and partially invariant to 3D projection. These features are efficiently detected through a staged filtering approach and are highly distinctive.

2.4.1 Model based face recognition

There are many popular real-world applications related to age synthesis [109] and estimation. Computer-aided age synthesis significantly relieves the burden of tedious manual work while at the same time providing more photo realistic effects and high-quality pictures. Age estimation by machine is useful in applications where one do not need to specifically identify an individual, but want to know his or her age.

Age estimation is a type of soft biometrics [110] that provides ancillary information of the users identity information. It can be used to compliment the primary biometric features, such as face [5], fingerprint [5], iris [5], and hand geometry [5], to improve the performance of a primary (hard) biometrics system. In real face recognition or identification applications, it is often the case that the system needs to recognize or identify faces after a gap of several years [15, 35, 50, 111], such as passport renewal and border security [20], which reveals the importance of age synthesis. Integrated with a dynamic ageing model, the face recognition or identification system can dynamically tune the model parameters by considering the shape or texture variations during the ageing process. System robustness to time gap can be significantly improved [112].

From a technical point of view, males and females may have different face ageing patterns displayed in images due to the different extent in using makeups and accessories [112, 113]. Many female face images may potentially show younger appearances. There are still two open problems remaining on how to:

1. Build personalized age synthesis models and

2. Extract general discriminative features for age estimation, reducing the negative influence of individual differences.

Moreover, the burden of acquiring large-scale databases which cover enough age range with chronological aged face images makes the estimation tasks more difficult to achieve. Although web image mining can help the data collection [114], it is usually hard or even impractical to collect a large database of large amount of subjects who can provide a series of personal images in different ages.

2.5 Ageing database

The availability of two publicly available ageing databases MORPH-II [15] (Figure 2.5) and FG-NET-AD [115] (Figure 2.6) played an important role in initiating an increased interest in research related to facial ageing among the computer vision community during the last decade.

Figure 2.5 Example of MORPH-II ageing database subject [7]

Figure 2.6 Example of FG-NET ageing database subject [8]

Some example of the public availability of ageing datasets are as follows:

• MORPH Face Ageing Database [15]: The publicly available MORPH face database was collected by the Face Ageing Group at the University of North Carolina at Wilmington for the purpose of face biometrics applications. This longitudinal database records individuals metadata, such as age, gender, ethnicity, height, weight, and ancestry, which is organized into two albums. Album I contains 1,724 face images of 515 subjects taken between 1962 and 1998. The ages range from the average of 27.3 to maximum 68 years. There are 294 images of females and 1,430 images of males. The age span is from 46 days to 29 years. Album II original database consists of 55,132 facial images acquired from 13,618 subjects. Population age range is 16-67 when first enrolled although most of the images are from 20-49 year old. Data was acquired over a period of 5 years but not everybody provided samples every year.

- FG-NET Ageing Database [116]: The FG-NET ageing database is publicly available. It contains 1,002 high-resolution color or gray-scale face images of 82 multiple-race subjects with large variation of lighting, pose, and expression. The age range is from 0 to 69 years with chronological ageing images available for each subject (on average, 12 images per subject).

- Gallagher and Chen's Web-Collected Database [117, 118]: The Web-collected database reported in [117] contains 28,231 faces (in 5,080 images) from the Flickr.com image search engine using three-group searches such as "wedding+ bride+ groom+portrait" "group shot" or "group photo" or "group portrait" "family portrait" . The age labels are of seven categories: 0-2, 3-7, 8-12, 13-19, 20-36, 37-65, and 66+. The median face has only 18.5 pixels between the eye centers, and 25 percent of the faces have under 12.5 pixels [118].

The FG-NET Ageing database, MORPH database, and Gallaghers Web-collected database are all publicly available. The other databases are potentially available by contacting the owners as shown in Table 2.3. The MORPH, YGA, LHI, Nis, and Gallaghers Web-collected databases are large-size ageing face databases which are promising for data-driven statistical algorithms of age estimation, such as AAM and age manifold with the regression model. The FG-NET is a baseline database for comparisons with many existing age synthesis and estimation techniques, such as AGES. The AI&R, Iranian, and LHI databases contain relatively high resolution 2D images which can be effectively used for age synthesis evaluation. The 3D morphable database is specifically collected for 3D age synthesis. The other databases mentioned here were not widely used but might be useful for some specific applications.

Table 2.3 A summary of ageing face datasets

Name	Number of faces	Number of subjects	Age range
MORPH-I [15]	1690	515	15-68
MORPH-II[15]	55134	13000	16-99
FG-NET [45]	1002	82	0-69
Burts Caucasian Face Database [45]	147	147	20-62
WIT-DB / Waseda [119]	26222	5320	3-85
Lotus Hill Research Institute [15, 120]	50000	50000	9-89
Human and Object Interaction Processing [121]	306600	300	15-64
AI&R Asian Face Database [122]	34	17	22-61
Gallagher and Chen's Web-Collected Database [117, 118]	28231	28231	0-66+
Yamaha Gender and Age Database [112, 123]	8000	1600	0-93
Ni's Web-Collected Database [114]	219892	77021	1-80
3D Morphable Database [124, 125]	438	438	8-16
Face Recognition Grand Challenge [126]	44278	568	Teenagers and adults
Internet Ageing Image Database [114]	219892	219892	1-80

2.6 Conclusion

Face ageing is a progressive accumulation of changes with time, and how fast a person age varies from one individual to another. Ageing effects both shape and texture, and is usually contributed by our genes, environmental influences and lifestyle. An important consequence of the ageing process is that enrolled templates become unrepresentative of the input (user) data after a certain time lapse as a result of change in the data distribution of an individual.

Existing age progression algorithms are capable of generating face images with ageing effects superimposed. Age progression algorithms are not yet to a stage to produce highly accurate predictions for different subjects mainly due to the diversity of ageing effects, the dependence of the ageing process on external factors/events that may occur in a subjects life and also due to the compounded effect of ageing variation with other within-subject variations (i.e. expression, pose and illumination).

The next chapter represent experimental methodology and investigation results will be presented.

CHAPTER 3

Experimental Methodology

3.1 Introduction

This chapter provides the details of the experimental framework developed for the investigation and evaluation of the proposed approach. This chapter also elaborates the different types of lateral and longitudinal ageing scenarios and the performance measures that are used to assess the verification rates. The implementation of the system as used for training and testing is also explained in this chapter. Furthermore, the database that is developed to investigate the age sensitivity of face algorithms is also covered in this chapter. Face ageing is still relatively a green area of research, nevertheless, there are some databases that are available publicly that has been listed in more details in Section 2.5. The MORPH-II database [15] has been used in this study.

The remainder of the chapter is organised as follow. Section 3.2 introduces the proposed system and presents the experiments scenarios. Section 3.3 provides details of the system implementation and test setup. It consists of database, the selected subjects, lateral face ageing and longitudinal template ageing using linear and non-linear mapping approaches. Section 3.4 explains the types of feature extraction approaches that are used in the study in lateral and longitudinal case. Section 3.5 looks into details on the four types of projection algorithms in the lateral ageing implementation. Section 3.6 relates the five types of distance metrics that was implemented in part of the study. Section 3.8 touches on the objective verification methods and metrics used in this study and finally Section 3.10 wraps up the contents of this chapter.

3.2 The proposed face recognition system

A typical face recognition system as shown in Figure 3.1 comprises of four modules such as image acquisition, preprocessing, feature extraction, and matching. The matching stage compares the acquired test image with relevant ones in the template database and generates the decision. The database module retains the enrolled templates to facilitate the matching process.

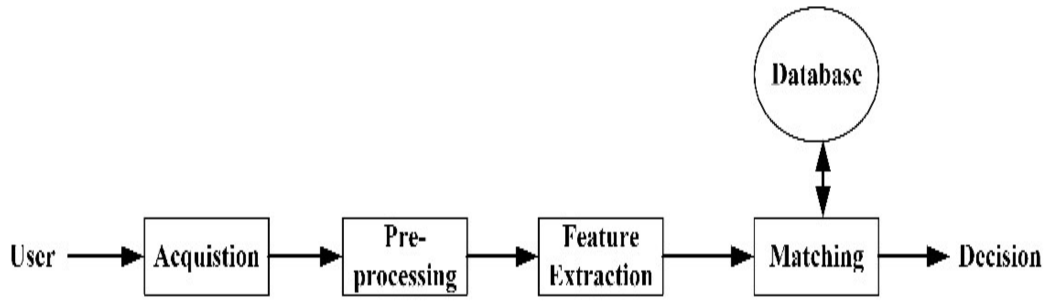

Figure 3.1 Basic modules of a face recognition system

In the proposed system, an additional process is introduced to stimulate aged the templates prior to comparison by the matching as shown in Figure 3.2. The motivation is to reduce the variation between the enrolled and test data in order to achieve a more reliable matching outcome.

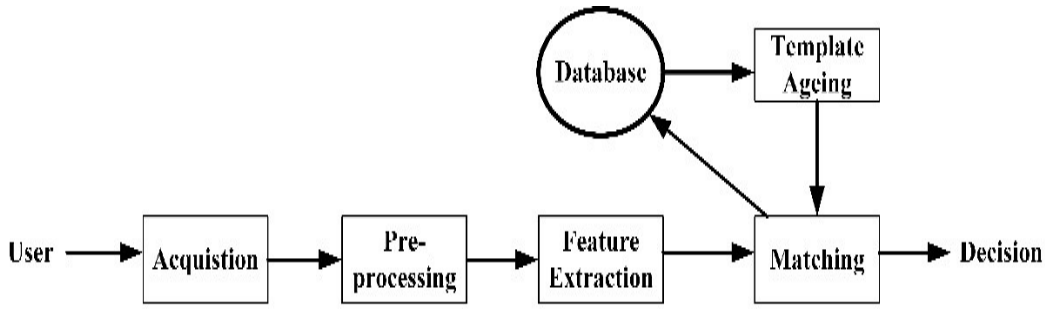

Figure 3.2 Proposed face recognition system with template ageing

3.2.1 Experiment scenarios

There are several phases of experiments in this study. Various types of template matching scenarios are first investigated to establish the benchmark performance for this study. The next scenario investigates the performance degradation of facial recognition systems due to the influence of age. The subsequent scenario emphasises on the performance degradation in facial recognition systems due to the age interval between the enrolled and the test samples, which in return, will propose a template ageing transformation as a remedy.

3.3 Implementation

The following section will describe the implementation of the proposed system. It will stretch on the details of the image pre-processing and normalisation as well as databases for the lateral and longitudinal face ageing scenarios.

3.3.1 Image pre-processing and normalisation

Before the feature extraction and matching, input images should be adapted and enhanced to the face recognition system, both in the training and/or the test stages.

Image pre-processing operations are necessary [127] for a good performance. First, the original RGB face images have to be converted to the grayscale form. Then, some contrast and illumination adjustment operations will have to be performed so that the effect of illumination and contrast are reduced in the processed images.

The following steps describe image preprocessing stages as shown in Figure 3.3

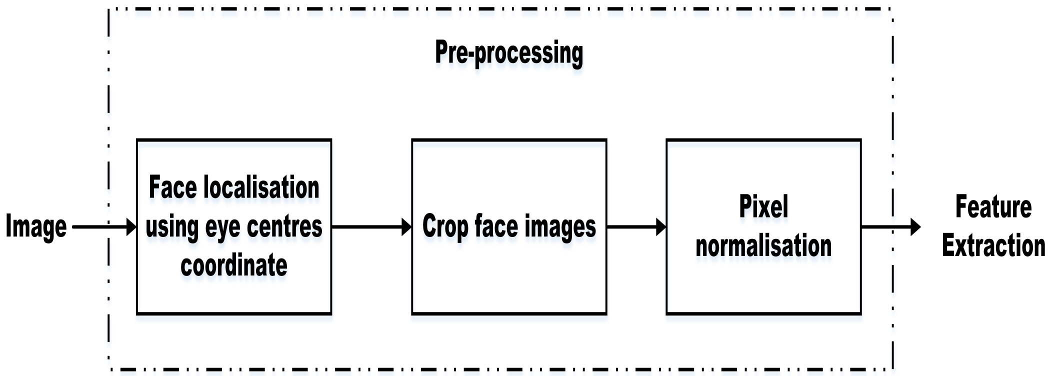

Figure 3.3 Face Recognition - Preprocessing Block Diagram

The location of the eye centres are detected manually. The image is then scaled and rotated so that the distance between eye centres are the same and the tilted head adjusted in such a way that the eyes lie horizontal to the image. The resulting images will then be cropped to 200×260 pixels. The recommended dimension by NIST (National Institute of Standards and Technology) is $width \times width/0.75$ [128] in such a way that faces are centred in the cropped image. Finally the pixel intensities are normalized. Normalization include rotating the image to make it align so that the tilted image is approximately around 0 degrees. Figure 3.4 presents the preprocessing results as the RGB image is cropped and normalised.

Figure 3.4 (a) MORPH-II image is 400 x 480
(b) Recommended dimension by NIST
(c) Image cropped to 200 x 260
(d) Normalized and rotated

3.4 Feature extraction approach

Face feature extraction represents the third part of the verification/identification process [91]. Two types of feature extraction approaches are explored in this study and will be explained in this section. In the lateral face ageing study, only the Gabor filter-based face feature extraction were investigated whereas for the proposed template ageing system (in Section 3.2) two feature extraction methods; Gabor filter-based and LBP will be investigated for both linear and non-linear mappings methods.

3.4.1 Gabor filter-based

This section briefly reviews the basic principles of face recognition using Gabor filters. It commences by introducing the Gabor filters and the Gabor face representation and proceeds by highlighting some characteristics of the filters, which affect the Gabor face representation and consequently the recognition performance of Gabor filter based recognition techniques.

In this study, two-dimensional Gabor filtering were chosen for feature extraction.

The principal motivation to use Gabor filters is the biological relevance that the receptive field profiles of neurons in the primary visual cortex of mammals are oriented and have characteristic spatial frequencies [9].

3.4.2 Using Gabor filter for face recognition

Gabor filters are bandpass filters which are used in image processing for feature extraction, texture analysis [129], and stereo disparity estimation [130, 131]. By extending these functions to two dimensions it is possible to create filters which are selective for orientation. Under certain conditions the phase of the response of Gabor filters is approximately linear. This property is exploited by stereo approaches which use the phase-difference of the left and right filter responses to estimate the disparity in the stereo images [132, 133].

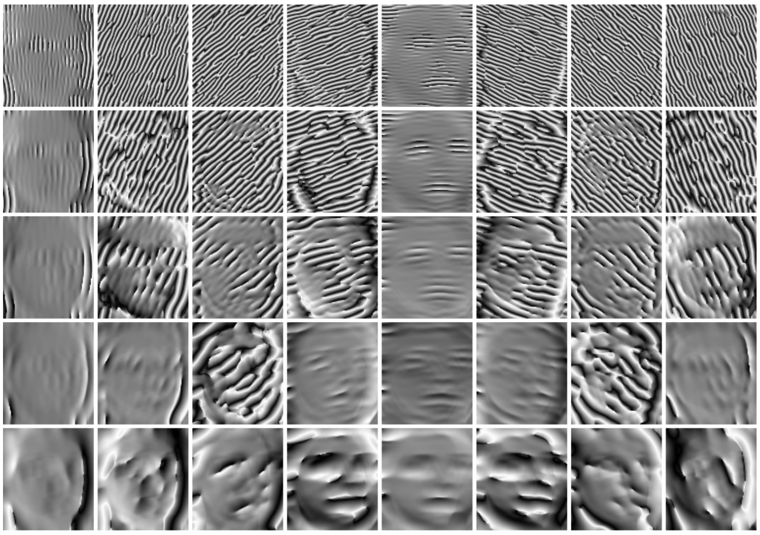

Figure 3.5 Example of gabor filter

Figure 3.5 shows an example of how Gabor filter works where as Figure 3.6 shows a visual comparison between the real part of the Gabor filter bank Γ and the real part of the eigenvector of all r correlation matrices Σ_v.

While the orthogonality of the principal Gabor filters might have a positive effect on the (principal) Gabor face representation in terms of compactness, this comes at

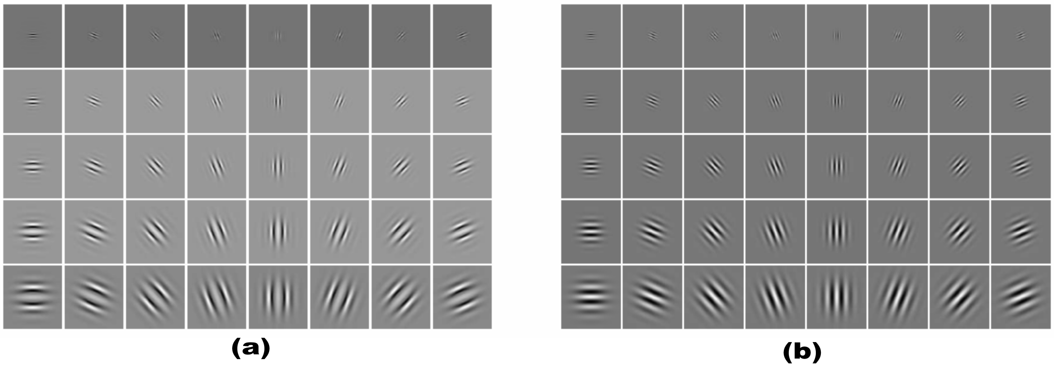

Figure 3.6 Visual comparison of (a) the real part of the eigenvectors of the correlation matrices and (b) the real part of the Gabor filter bank [9]

a price, as the filters are not localized optimally in the spatial nor in the frequency domain anymore. This fact is illustrated in Figure 3.7, where the top row shows the magnitudes of the classical Gabor filters at different scales and the bottom row shows the magnitudes of the principal filters at different scales.

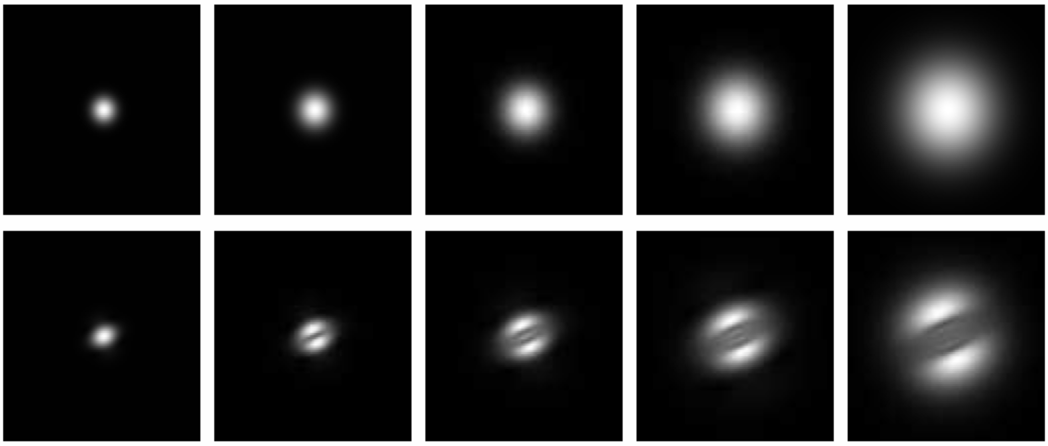

Figure 3.7 The magnitude of the Gabor kernels at different scales. The kernels exhibit desirable characteristics of spatial frequency, spatial locality, and orientation selectivity

Gabor filters can exploit salient visual properties such as spatial localization, orientation selectivity, and spatial frequency characteristics [51, 91]. Considering these overwhelming capacities and its great success in face recognition, this thesis will explore Gabor features to represent the face image and produces verification task.

The amount of data (in the Gabor face representation) is then commonly reduced to a more manageable size by exploiting various downsampling, feature selection and subspace projection techniques before it is finally fed to a classifier [134]

A set of 40 Gabor filters with different frequencies and orientations were used for the feature extraction here. The feature vectors, consists of the Gabor magnitudes, are extracted for all the experiments in this study.

In the longitudinal ageing for non-linear mapping transformation method, neural network is being introduced using Gabor filter. A boosting learning process is used to reduce the feature dimensions and make the Gabor feature extraction process substantially more efficient [135]. Combining optimized Gabor features with Neural Networks [103] reduces computation and memory cost of the feature extraction process, but also achieves very accurate recognition performance. Actually, training process in a neural network does not consist of a single call to a training function. Instead, the network was trained several times on various noisy images [136].

In GaborPCA algorithm, feature vectors are generated at the feature points as a composition of Gabor wavelet transform coefficients. Here k^{th} feature vector of he i^{th} reference face and R is the response of the face image (to the Gabor filter) is defined as,

$$v_{i,k} = \{x_k, y_k, R_{i,k}\left(x_k, y_k\right)\} \tag{3.1}$$

After feature vectors are constructed from the test image, they are compared to the feature vectors of each reference image in the database.

Generally, it is difficult to deal with a high dimensional image space. So this Gabor wavelet method is used to reduce the space dimension by down sampling each $G(u, v)$ and concatenating its rows to form a 1D feature vector is proposed and used extensively. This algorithm is tested for the task of verification using neural network classifier. To reduce the dimensionality of the vector space and obtain more useful features for subsequent pattern discrimination and associative recall, the PCA technique is used here. The results clearly show that Gabor filters improve the performance of raw image data with the operating points corresponding to high False Accept Rate (FAR).

3.4.3 Local Binary Pattern (LBP)

The second feature extraction method explored is the Local Binary Pattern (LBP). The local binary pattern operator is an image operator which transforms an image into an array or image of integer labels describing small-scale appearance of the image. These labels or their statistics, most commonly the histogram, are then used for further image analysis.

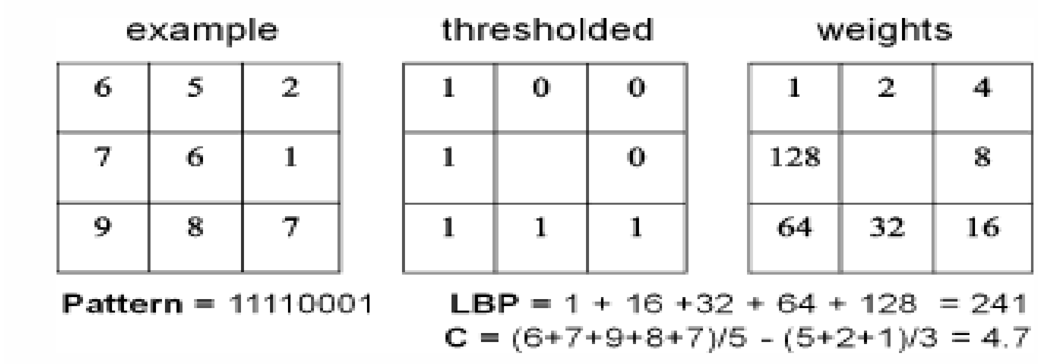

Figure 3.8 The original LBP algorithm

The basic local binary pattern operator, introduced by Ojala et al. [137], was based on the assumption that texture has locally two complementary aspects, a pattern and its strength. In that work, the LBP was proposed as a two-level version of the texture unit to describe the local textural patterns [138].

The original version of the local binary pattern operator works in a 3×3 pixel block of an image. The pixels in this block are threshold by its center pixel value, multiplied by powers of two and then summed to obtain a label for the center pixel.

As the neighborhood consists of 8 pixels, a total of $2^8 = 256$ different labels can be obtained depending on the relative gray values of the center and the pixels in the neighborhood. See Figure 3.8 for an illustration of the basic LBP operator. An example of an LBP image and histogram are shown in Figure 3.9

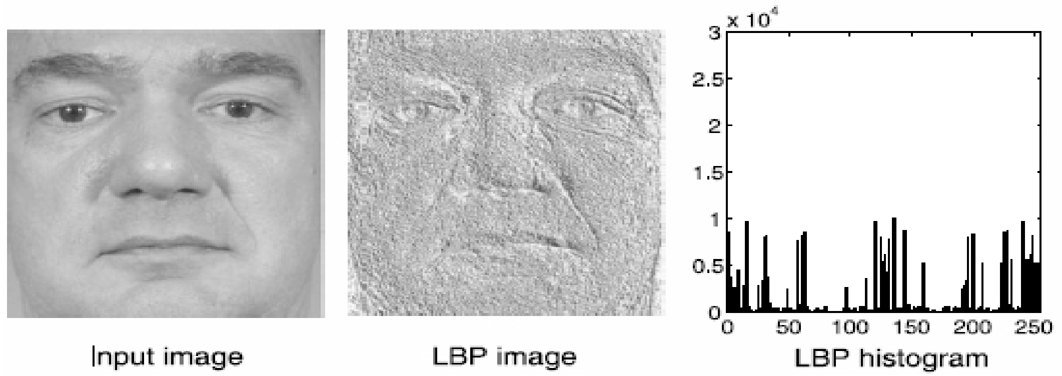

Figure 3.9 Example of an input image, the corresponding LBP image and histogram [10]

An extension to the original operator uses so called uniform patterns [139]. For this, a uniformity measure of a pattern is used: U (pattern) is the number of bitwise transitions from 0 to 1 or vice versa when the bit pattern is considered circular.

A local binary pattern is called uniform if its uniformity measure is at most 2. For example, the patterns 00000000 (0 transitions), 01110000 (2 transitions) and 11001111 (2 transitions) are uniform whereas the patterns 11001001 (4 transitions) and 01010011 (6 transitions) are not.

The reasons for omitting the non-uniform patterns are twofold. First, most of the local binary patterns in natural images are uniform. Ojala et al. [139] noticed that in their experiments with texture images, uniform patterns account for a bit less than 90% of all patterns when using the (8, 1) neighborhood and for around 70% in the (16, 2) neighborhood. In experiments with facial images it was found that 90.6% of the patterns in the (8, 1) neighborhood and 85.2% of the patterns in the (8, 2) neighborhood are uniform [39].

The second reason for considering uniform patterns is the statistical robustness. Using uniform patterns instead of all the possible patterns has produced better recognition results in many applications. On one hand, there are indications that uniform patterns themselves are more stable, i.e. less prone to noise and on the other hand, considering only uniform patterns makes the number of possible LBP labels significantly lower and reliable estimation of their distribution requires fewer samples. Table

Table 3.1 Example of uniform and non uniform Local Binary Patterns

Circular Binary Pattern	No of bitwise transitions	Uniform pattern
11111111	0	Yes
00001111	1	Yes
01110000	2	Yes
11001110	3	No
11001001	4	No

3.1 presents an example of types LBP.

The uniform patterns allows to see the LBP method as a unifying approach to the traditionally divergent statistical and structural models of texture analysis [140]. Each pixel is labeled with the code of the texture primitive that best matches the local neighborhood. Thus each LBP code can be regarded as a micro-texton. Local primitives detected by the LBP include spots, flat areas, edges, edge ends, curves and so on. Some examples are shown in Figure 3.10 with the $LBP_{8,R}$ operator.

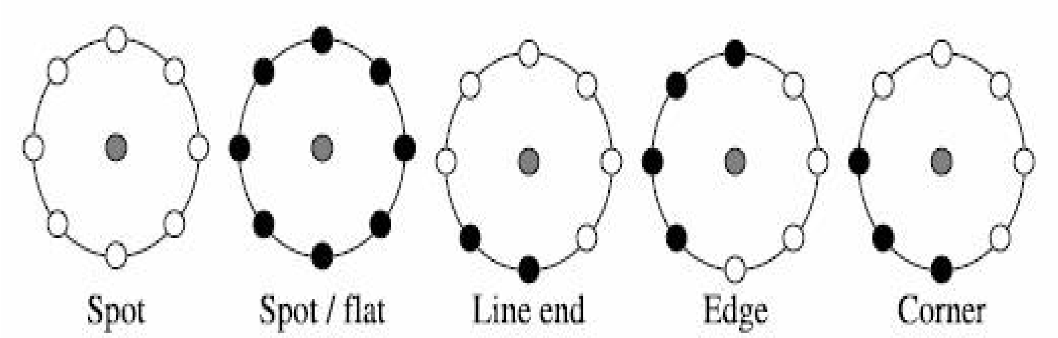

Figure 3.10 Different texture primitives detected by the LBP

In Figure 3.10, ones are represented as black circles, and zeros are white. The combination of the structural and statistical approaches stems from the fact that the distribution of micro-textons can be seen as statistical placement rules. The LBP distribution therefore has both of the properties of a structural analysis method: texture primitives and placement rules. On the other hand, the distribution is just a statistic of a non-linearly filtered image, clearly making the method a statistical one. For these reasons, the LBP distribution can be successfully used in recognizing a wide variety of different textures, to which statistical and structural methods have normally been applied separately.

3.4.4 Using LBP for face recognition

In 2004, a novel facial representation for face recognition based on LBP features was proposed [141]. In this approach, the face image is divided into several regions from which the LBP features are extracted and concatenated into an enhanced feature vector to be used as a face descriptor [39, 141]. This approach has evolved to be a growing success. It has been adopted and further developed by a large number of research groups and companies around the world. The approach and its variants have been used in problems such as face recognition and authentication, face detection, facial expression recognition, gender classification and age estimation.

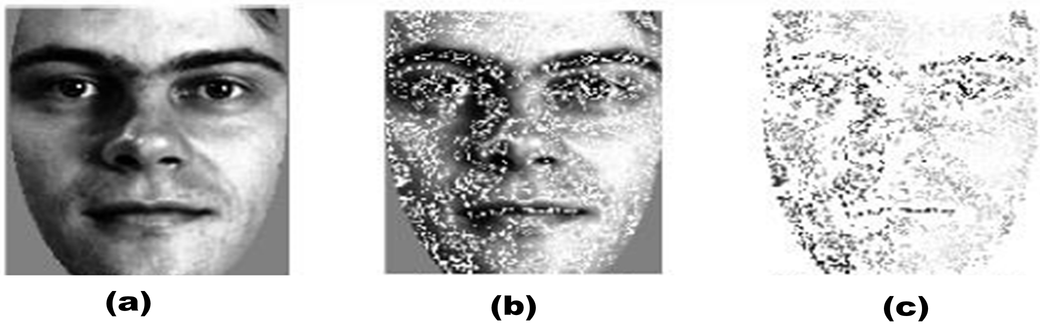

Figure 3.11 (a) Face image split in an image with only pixels with uniform patterns and (b) in an image with only non-uniform patterns, (c) by using $LBP^{u2}_{16,2}$ [10]

Figure 3.11 shows an example of an image which is split in an image with only pixels with uniform patterns and in an image with only non-uniform patterns. These images are created by using the $LBP^{u2}_{16,2}$ operator. It occurs that the image with only pixels with uniform patterns still contains a considerable amount of pixels. So, 79% of the pixels of the image have uniform patterns (with $LBP^{u2}_{8,2}$ this is even 86%). Another striking thing is the fact that, by taking only the pixels with uniform patterns, the background is also preserved. This is because the background pixels all have the same color (same gray value) and thus their patterns contain zero transitions. It also seems that much of the pixels around the mouth, the noise and the eyes (especially the eyebrows) have uniform patterns.

Once the LBP for every pixel is calculated, the feature vector of the image can be constructed. For an efficient representation of the face, first the image is divided into k^2 regions. In Figure 3.12 a face image is divided into $7^2 = 49$ regions. For every region a histogram with all possible labels is constructed. This means that every bin in a histogram represents a pattern and contains the number of its appearance in the region. The feature vector is then constructed by concatenating the regional histograms.

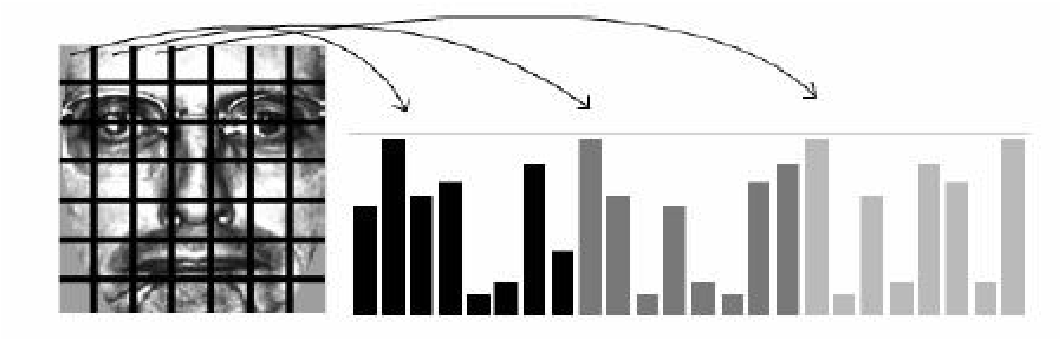

Figure 3.12 Face image divided into 49 regions, with for every region a histogram [10]

For the LBP algorithm, feature vectors are constructed once the LBP for every pixel is calculated. For an efficient representation of the face, first the image is divided into k^2 regions. In constructing the feature vector, a small area on the borders of the image is not used. For an $N \times M$ image, the feature vector is constructed by calculating the LBP code for every pixel (x_c, y_c) with $x_c \in \{R+1, ..., N-R\}$ and $y_c \in \{R+1, ..., M-R\}$ with $k \times k$ regions, then the histogram for region (k_x, k_y), with $k_x \in \{1, ..., k\}$ and $k_y \in \{1, ..., k\}$, where P is the sampling points and radius R can be defined as:

$$H_i(k_x, k_y) = \sum_{x,y} I\{LBP_{P,R}(x, y)\} = i = 1, ..., P(P+1) + 3) \tag{3.2}$$

$$x \in \begin{cases} \{R+1, ..., N/k\} & where \quad k_x = 1 \\ \{(k_x - 1)(N/k) + 1, ..., N - R\} & where \quad k_x = k \\ \{(k_x - 1)(N/k) + 1, ..., k_x(N/k)\} & else \end{cases}$$

$$y \in \begin{cases} \{R+1, ..., M/k\} & where \quad k_y = 1 \\ \{(k_y - 1)(M/k) + 1, ..., M - R\} & where \quad k_y = k \\ \{(k_y - 1)(M/k) + 1, ..., k_y(M/k)\} & else \end{cases}$$

in which L is the label of bin i and $I(A) = \begin{cases} 1, & A \ is \ true \\ 0, & A \ is \ false \end{cases}$

The feature vector is effectively a description of the face on three different levels of locality: the label contain information about the patterns on a pixel-level, the regions, in which the different labels are summed, contain information on a small regional level and the histograms give a global description of the face.

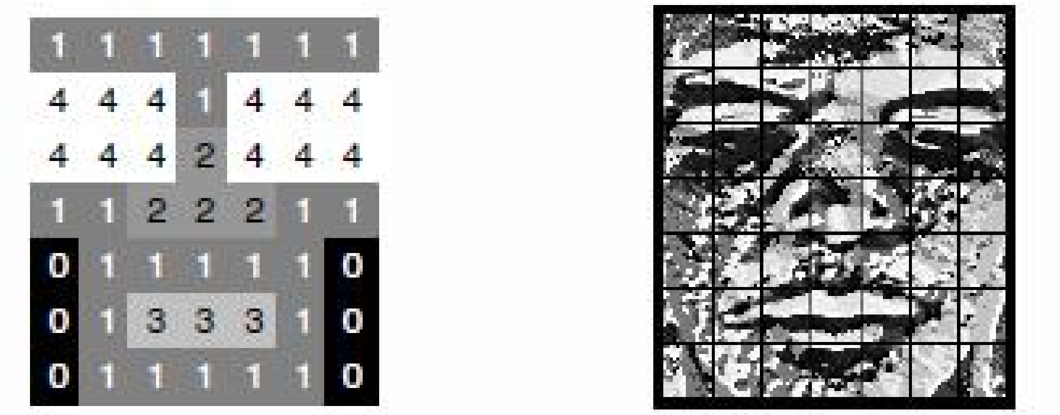

Figure 3.13 Face regions weights sample taking into account human recognition system [11]

Dividing the face image in smaller regions can help to emphasize facial features with essential information for face description and to remove irrelevant information by using an adequate weight value. Obviously, nose, mouth and eyes regions are likely to provide more important information in face description task, than bottom left and for example, the right blocks where normally background areas, hair and even cloth can be present.

Moreover, neurophysiologic studies based on face recognition human system evidence that most important face features are eyes and eyebrows, followed by mouth and nose (see [142, 143]). Therefore, and taking into account human system knowledge, Figure 3.13 shows a possible good choice for weighting face areas in face recognition procedure.

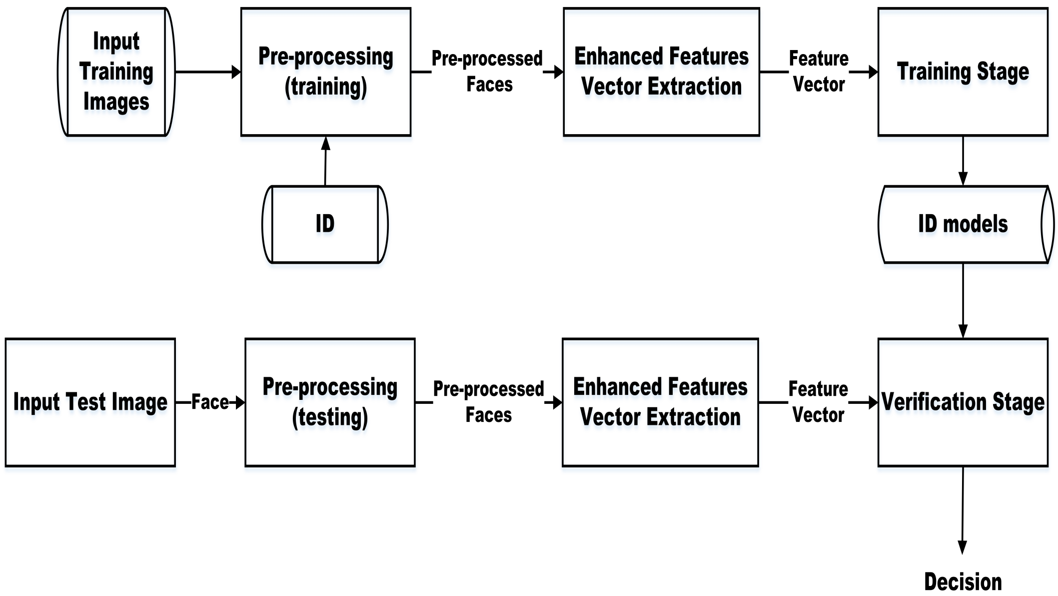

Figure 3.14 Generic face recognition scheme

In face verification step, given a new input face image, the 2891 features vector is extracted and compared to all available models by means of a dissimilarly metric. The input face will be identified with the minimum ID model distance. Following Figure 3.14 introduces the generic face recognition scheme proposed for this study.

- Training stage: input face samples are introduced in the pre-processing block together with an ID label to crop faces from input image and resize them. Then $LBP_{8,1}^{u2}$ enhanced feature vector is extracted and used to feed the training block that will output a mean value model for each ID in the training database.

- Test stage: for each new test image, it will go through the same pre-processing stage for image quality improvement. Then $LBP_{8,1}^{u2}$ enhanced feature vector will be extracted and compared with dissimilarity distance matrix (Chi-Square distance).

For the experiments in this study, the face regions with mainly background pixels are left out by assigning them zero weights as shown in Figure 3.15. The eye and mouth-regions are assigned with the highest weight, because in the numerous trial and testing, it is found that the eyes and mouth reveal important characteristic in the ageing for a person.

0	0	1	1	1	0	0
0	1	1	1	1	1	0
1	2	2	1	2	2	1
1	4	4	3	4	4	1
1	1	2	2	2	1	1
0	1	2	2	2	1	0
0	0	1	1	1	0	0

Figure 3.15 Region weights used in this experiment

The face image is divided into 7×7 blocks and concatenated the block histograms into a single vector (7×7 blocks $\times$ 59-bin/block $= 2891$ features). These features are calculated for each individual sample and a model is computed for each ID subset.

3.5 Feature Projection Techniques

For limited training samples, it is often difficult to model class distributions well if the feature vectors are too large. Use of a feature reduction scheme can reduce the complexity of the task and hence improve the performance of a system. Many techniques have been proposed where high dimensional feature vectors can be projected to a lower dimensional space while minimally compromising (rather often improving) attributes such as information content, ability to discriminate, etc. Four such schemes have been investigated in this study as described below.

3.5.1 Principal Component Analysis (PCA)

PCA [144] is a probabilistic method for finding patterns in data with high dimensions. PCA is an orthogonal linear transformation that transforms the original data to a new coordinate system with fewer dimensions. This is done by calculating the eigenvalues and eigenvectors of the covariance matrix of the original data set. The covariance matrix contains values that indicate how much the dimensions vary from the mean with respect to each other. The eigenvectors with the highest eigenvalues contain the most information and are the principal components of the data set. By choosing only a few eigenvectors (with the highest eigenvalues) and multiplying these vectors with the original data set, a new data set is obtained with less dimensions, but still with most of the important characteristics of the original data.

In face recognition systems PCA can be used to reduce the dimensions of the original feature vectors of the face images [145]. By choosing only the most important eigenvectors, irrelevant information (e.g., pixels which represent background) are filtered out because of the low variance. Along the new axes the differences and similarities between a new image and the original images from the database are measured. Figure 3.16 shows the PCA approach using a flow chart.

3.5.2 Linear Discriminant Analysis(LDA)

LDA [146] is also a statistical method. It can be used to classify (new) objects (for example people) into two or more groups based on a set of features that describe the objects.

The main idea is that, after the linear transformation, the differences within the group are minimized and the differences between the groups are maximized. In the ideal case, a projection (a discriminant) is found that completely separates the different groups. A new object can then be assigned the the group with the highest conditional

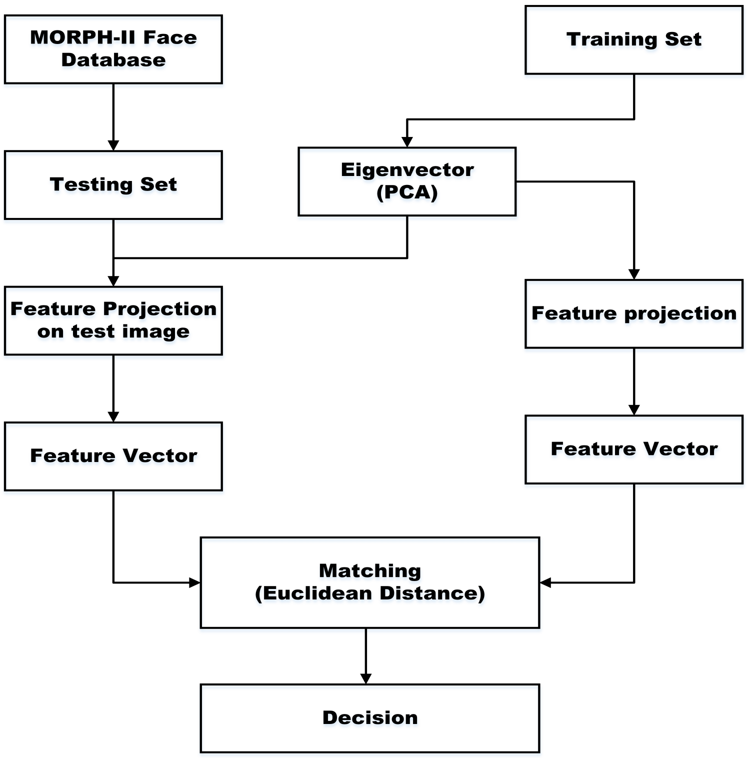

Figure 3.16 PCA approach for face recognition

probability. To improve the classification process, a linear transformation on the original feature set is performed to obtain a new and reduced set of features that describes a group in the best way.

LDA seeks to find directions along which the classes are best separated. That takes into consideration the scatter within-classes but also the scatter between-classes. The eigenvectors of LDA are called fisherfaces. LDA transformation dependence on number of classes, number of samples, and original space dimensionality. LDA derives a low dimensional representation of a high dimensional face feature vector space. The discriminating feature vector is given by the coefficients of the covariance matrix for the LDA method. The transformation matrix projects the face vector. The projection coefficients are used for the feature representation of each face image. In the proposed scheme the column vectors of the matrix are referred as fisherfaces.

3.5.3 Kernel Principal Component Analysis (KPCA)

PCA [147] has been widely used in many pattern recognition applications, such as face recognition [51, 148] and remote sensing [149]. Unlike the traditional PCA approach, where only linear projection is performed to seek a best mapping of the original dataset, the kernel PCA [150] relaxes the linear constraint, and allows arbitrary high-order projections among the input data through smaller dimension to higher dimensional mappings. Kernel PCA linearly represents the non-linear problem by means of mapping the low-dimensional input space, which is usually non-linear separable, into a linear separable high-dimensional feature space. More specifically, the underlying principle of kernel PCA is addressed by Covers theorem [151, 152]. The low to high dimensional mapping is defined implicitly by a so-called kernel function, which efficiently computes the inner product as a direct function of the input space. Without explicitly computing the mapping function, the kernel PCA becomes more computationally feasible [153].

Gabor-based kernel PCA [154] is combined with the Gabor wavelet decomposition of the sample dataset and kernel PCA for pattern recognition. First, Gabor wavelet decomposition is applied to the sample dataset to obtain the Gabor features of the input data. Then the Gabor feature vector, “W”, is fed into the kernel PCA algorithm.

In other words, the Gabor feature space is regarded as the input space of the kernel PCA. Through the kernel PCA, Gabor feature space is mapped to a high-dimensional feature space, “C”, making the high-dimensional features linearly separable by PCA in that space. Finally,the nearest neighbor classifier is used in the high-dimensional space, “C”, to differentiate between individuals using various distance metric.

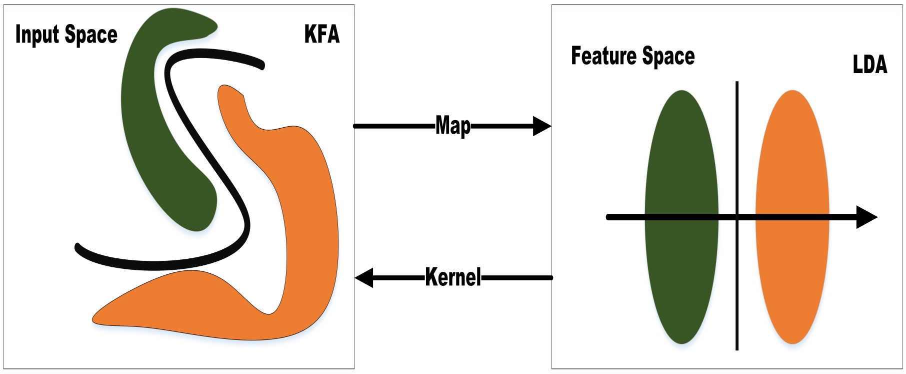

Figure 3.17 Kernel Fisher Analysis

3.5.4 Kernel Fisher Analysis (KFA)

The principle of KFA can be illustrated in Figure 3.17 owing to the severe non-linearity, it is difficult to directly compute the discriminating features between the two classes of patterns in the original input space (left). By defining a non-linear mapping from the input space to a high-dimensional feature space (right), a linearly separable distribution can be obtained in the feature space. Then LDA, the linear technique, can be performed in the feature space to extract the most significant discriminating features. However, the computation may be problematic or even impossible in the feature space owing to the high dimensionality. By introducing a kernel function which corresponds to the non-linear mapping, all the computation can conveniently be carried out in the input space. The problem can be finally solved as an eigen-decomposition problem like PCA, LDA and KPCA.

3.6 Distance Metrics

Identity of the user are verified by comparing the transformed feature vector from a live sample with the ones already stored in the template database. Distances between two vectors can be treated as a measure of dissimilarity of the two biometric samples.

Various distance metrics can be found in the literature of which five metrics have been chosen for this study. Mathematical description of these metrics are as follows:

3.6.1 Euclidean Distance

It is the most commonly used metric based on the Pythagorean formula [155]. If $u = (u_1, u_2, \ldots, u_N)$ and $v = (v_1, v_2, \ldots, v_N)$ are two feature vectors, the Euclidean distance is given by

$$D_{euc}(u, v) = \sqrt{\sum_{N}^{i=1} (u_i - v_i)^2} \tag{3.3}$$

3.6.2 Manhattan / CityBlock Distance

More formally, Manhattan distance, also known as the L1-distance and CityBlock distance, between two points in an Euclidean space with fixed Cartesian coordinate system is defined as the sum of the lengths of the projections of the line segment between the points onto the coordinate axes [156].

For example, in the 2D plane, the Manhattan distance between the point P_1 with coordinates (u_1, v_1) and the point P_2 at (u_2, v_2) is shown in Equation 3.4

$$D_{ctb}(u, v) = \sum_{i=1}^{N} |u_i - v_i| \tag{3.4}$$

Notice that the Manhattan distance depends on the choice on the rotation of the coordinate system, but does not depend on the translation of the coordinate system or its reflection with respect to a coordinate axis.

3.6.3 Cosine Similarity

Cosine similarity is measure of similarity based on the cosine of the angle between two vectors [157]. Two vectors of same orientation have a cosine similarity value of 1 which reduces to -1 as the angle between the vectors increase to 180^o. The mathematical expression is shown in Equation 3.5 where a minus sign is introduced to convert the measure to dissimilarity metric in line with the other metrics here. Hence the cosine distance between two vector u and v is given by

$$D_{cos}(u,v) = -\frac{u.v}{||u||||v||} = -\frac{\sum_{i=1}^{N} u_i v_i}{\sqrt{\sum_{N}^{i=1} u_i^2} . \sqrt{\sum_{N}^{i=1} v_i^2}} \tag{3.5}$$

Cosine Similarity will generate a metric that says how related two objects are by looking at the angle instead of magnitude, as in the examples below in Figure 3.18:

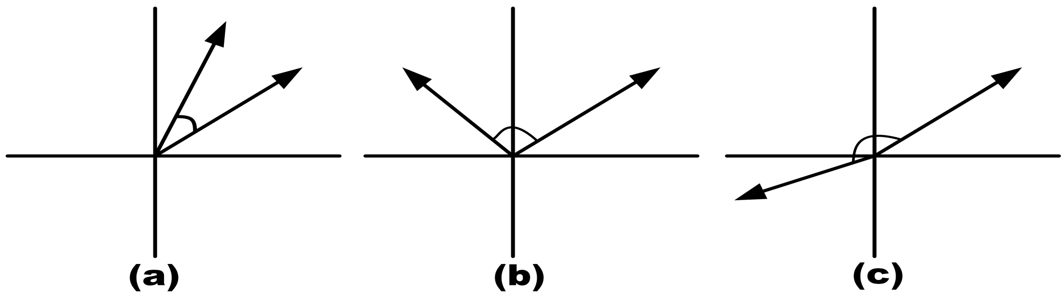

Figure 3.18 The Cosine Similarity values , (a) Score vectors in same direction. Cosine of angle is near 1, (b) Cosine vectors are nearly orthogonal. Cosine of angle is near 0 (90 deg.), (c) Score vectors in opposite direction. Cosine of angle is near -1

Note that even if a vector pointing to a point far from another vector, they still could have an small angle and that is the central point on the use of Cosine Similarity, the measurement tends to ignore the higher term count on objects.

3.6.4 Mahalanobis-Cosine Distance

It is similar to cosine distance but computed in the Mahalanobis space [158] as shown in Equation 3.6.

$$D_{MahCos}(u,v) = -\frac{m.n}{||m||||n||} where, m_i = \frac{u_i}{\sigma_i}, n_i = \frac{v_i}{\sigma_i} \tag{3.6}$$

Here σ is the standard deviation of the i^{th} dimension in the Mahanalobis space.

3.6.5 Chi-square Statistic

This distance metrics is applied for LBP method as this is the recommended metrics calculation for this feature. To compare a new input sample u with a template v, the Chi-Square distance is defined Equation 5.6:

$$D\chi^2(u,v) = \sum_{j=1}^{k^2}\left(\sum_{i=1}^{P(P-1)+3}\frac{(u_{i,j}-v_{i,j})^2}{u_{i,j}+v_{i,j}}\right) \tag{3.7}$$

where in uniform LBP mapping there is a separate output label for each uniform pattern and all the non-uniform patterns are assigned to a single label. Thus, the number of different output labels for mapping for patterns of P bits is $P(P-1)+3$. For instance, the uniform mapping produces 59 output labels for neighborhoods of 8 sampling points, and 243 labels for neighborhoods of 16 sampling points. While k^2 stands for the number of block, in this case the face image is divided into 7×7 blocks and concatenated the block histograms into a single vector (7×7 blocks $\times$ 59-bin/block = 2891 features).

All these measures can be extended to the spatially enhanced histogram by summing up over each face region. Taking Chi-square statistic as the best option for face recognition as proposed by [159], it can be extended for each region and also a different weight can be given for each image block as shown in Equation 3.8.

$$D\chi^2(u,v) = \sum_{j=1}^{k^2} w_j \left(\sum_{i=1}^{P(P-1)+3} \frac{(u_{i,j} - v_{i,j})^2}{u_{i,j} + v_{i,j}} \right) \tag{3.8}$$

where w_j is the weight for region j.

3.7 Database

Data collection is crucial for the photorealistic age template matching. However, it is extremely difficult, in practice, to collect large scale ageing databases, especially when one wants to collect the chronometric image series from an individual.

The publicly available MORPH face database was collected by the Face Ageing Group at the University of North Carolina at Wilmington for the purpose of face biometrics applications. This longitudinal database records individuals' images as well as metadata, such as age, gender, ethnicity, height, weight, and ancestry, which is organized into two albums. Album I contains 1,724 face images of 515 subjects taken between 1962 and 1998. The ages range from 16-68 with the average of 27.3 year olds. There are 294 images of females and 1,430 images of males. The age span is from 46 days to 29 years. Album II database consists of 55,132 facial images acquired from 13,618 subjects. Population age range is 16-77 when first enrolled with a median age of 33 although most of the images are from 20-49 year olds. The average number of images per individual is 4 and the average time between photos is 164 days, with the minimum being 1 day and the maximum being 1681 days. Data was acquired over a period of 5 years but not everybody provided samples every year.

In this study, MORPH-II database [15] has been adopted to test the performance of the face recognition systems using various feature extraction, feature selection and distance metric combinations.

3.7.1 Lateral ageing subset

A subset of 100 images from the MORPH-II database [15] has been selected to test the performance of the face recognition system for lateral age effect.

For this study, the selected population is split into 5 age bands ("$\leq$ 19", "20-29", "30-39", "40-49", "$\geq$ 50") depending on their first enrollment age. 20 subjects were randomly picked for each group ensuring that there are equal gender distributions in each group. There are 6 images from each person. Of these 6 images, 3 were used for enrollment and the rest for verification.

3.7.2 Longitudinal ageing subset

Again, a subset of 80 images from the MORPH-II database [15] has been extracted to test the performance with the proposed template ageing method for the two face recognition systems.

For this study, this subset only included images where the time interval between enrollment and verification is 3 years. 80 subjects were picked with the age between 20-55 years old at their first enrollment. For example, subject 1 with enrollment age of 20 and three years later as the age of 23 are used for testing.

3.7.3 Soft biometrics subset

For longitudinal subjects, the populations are split into gender and age sub-groups for extended detailed experiments. There are 50 males and 30 females, while there are 45 people in the mature group (aged between 33-55 years old) and 35 in the young group (aged between 20-32 years old).

Table 3.2 shows the demographic breakdown of the dataset.

Table 3.2 Demographic breakdown of the dataset

Description	Mature	Young	
Male	29	21	50
Female	16	14	30
	45	35	80

3.8 Performance Analysis

For comparative performance analysis, the verification scenario has been implemented in this study. In a biometric verification scenario, FAR denotes the portion of the impostors accepted as genuine whereas FR) denotes the proportion of genuine users rejected as impostor. The FAR and FRR depend on the operational parameters. The phenomenon when FAR equals FRR, these errors are termed as EER. GAR is defined as $1 - FRR$.

ROC curve plots FRR versus FAR. Alternatively, ROC curve also plots GAR versus FAR. DET curve is similar to ROC curve except that the axes are often scaled non-linearly to highlight the region of error rates of interest. Commonly used scales include normal deviate scale and logarithmic scale.

3.9 Investigation Result

In this section, we look into both lateral and longitudinal face ageing using age group distribution to evaluate the effects of ageing on biometric system accuracy. We used a subset of the MORPH-II database. There are 100 subjects which consisted of equal gender distribution for this experiment. For both lateral and longitudinal ageing scenario, the images of these 100 subjects were acquired where the age intervals was between 0 - 5 years.

Table 3.3 shows a comparison of the initial experiments results on the reliability of the our system with performances reported for similar algorithms used for face recognition systems using ORL database. The performance of our approach can be seen to compare more favourable with the other methods considered and lends support to its applicability in face ageing scenario. A set of 40 Gabor filters, $5scales \times 8orientations$ were used as the baseline parameter since this is the most commonly used in previous face recognition studies which achieved good performance results [154, 160–163].

Table 3.3 Comparison of performance reports

Orientation vs Scales	Method	Author	Recognition Rate (in %)
8/5	Gabor + PCA + Euc	MageshKumar et al. [164]	87.0
8/5	Gabor + PCA + Euc	Our system	90.0

EER were calculated for a range of system parameters and are reported in this section. The preliminary experiment gathered all the 100 subjects together without any sub-grouping and calculated the EER different configuration. Table 3.4 summarises the overall EER for various combinations of selected algorithms with different distance metrics.

Table 3.4 Comparative EER (%) without sub-grouping

Feature Projection Algorithm	Distance Metric			
	Euclidean	CityBlock	Cosine	Mahalanobis-Cosine
PCA	12.9	20.2	13.6	14.2
LDA	8.1	8.7	4.0	4.3
KPCA	16.8	24.3	14.4	49.5
KFA	20.8	17.1	17.2	16.1

Looking into the PCA algorithm, the Euclidean distance produced the lowest EER of 12.9%. The Cosine and Mahalanobis Cosine distances, produced a comparable uniform EER with a difference of 0.6% . The worst distance combination for PCA algorithm is the CityBlock distance with EER of 20.2% which showed a difference of 7.3% from Euclidean distance. Meanwhile in the LDA algorithm, cosine and Mahalanobis Cosine distance performed at the best among the rest of the distance with EER of 4.0% and 4.3% respectively. KPCA algorithm has the worst performer in all distance combination.

Among the feature projection algorithms both PCA and LDA produced a better EER due to its linear technique when combined with the selected distance metrics. It can be observed that both non-linear technique i.e KPCA and KFA does not performed well when ageing is incorporated into the experiment. Among the distance metrics method, Cosine and Mahalanobis-cosine combinations performed slightly better than Euclidean and CityBlock as the former measures the similarity of vectors with respect to the origin, while the latter measures the distance between particular points

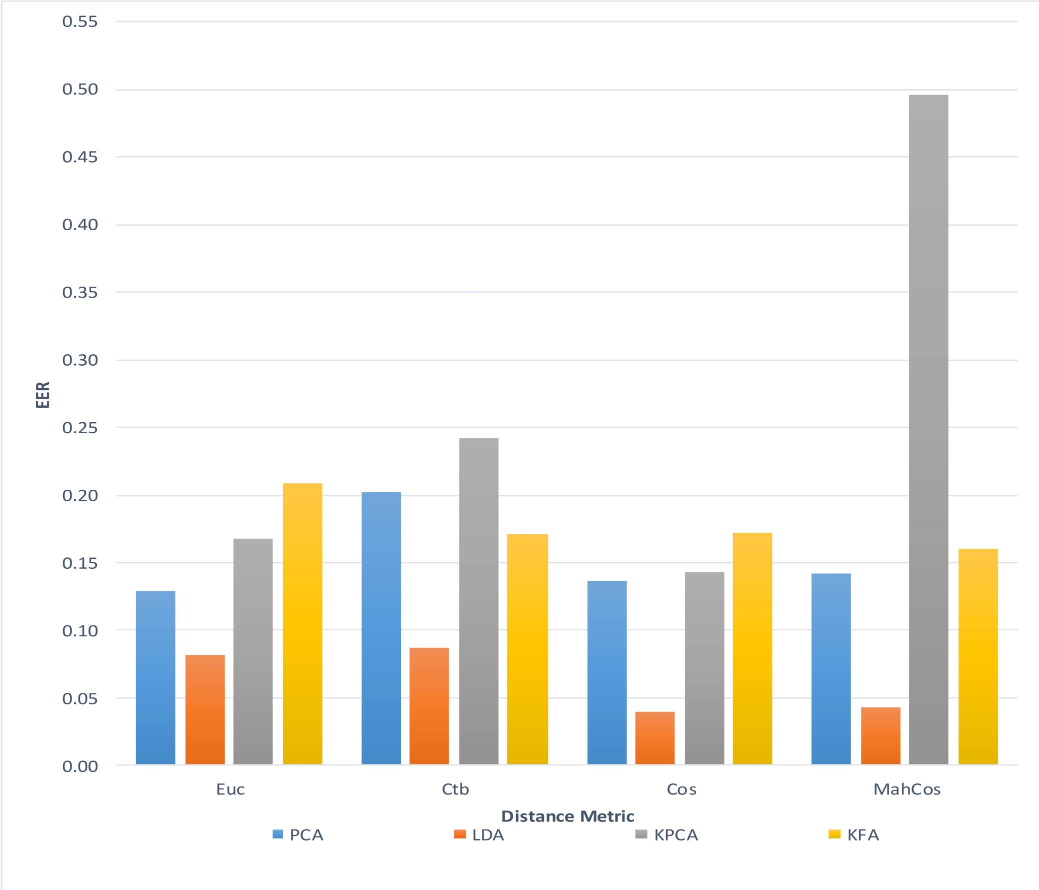

Figure 3.19 Effect of algorithm based on distance metric without sub-grouping

of interest along the vector.

From the Figure 3.19, it can be observed that LDA performed the best in all distance combinations with an average of 6.28% EER. If we look at distance metric, it can be readily observed that Cosine distance performed slightly lower than the other 3 projection algorithms. For PCA algorithm, the best metric combination belongs to Euclidean with EER of 12.9%.

Among the distance metrics, cosine distance almost always produced the lowest EER except when combined with KFA. Even then the margin is only about 1%. Mahalanobis Cosine distance also produced very low EER except when combined with KPCA. The best EER of 4.0% was achieved with the LDA-Cosine Distance combination. It is therefore evident that the choice of projection algorithm and distance metric has a significant impact on the achievable EER from a face recognition system.

3.9.1 Effect of ageing for different intervals

In this set of experiment, the same subset of 100 subjects are selected from the MORPH-II database. The age range in this subset is between "17 to 65" years old. For comparative performance analysis, the verification scenario has been implemented in this study.

Here, the subjects are divided into groups with age intervals of $0-2$, $0-3$ and $0-4$ to show the effect of face ageing when age intervals is applies to the face recognition algorithm. The "0" stands for enrollment age where "2", "3" and "4" represent the test samples of the same subjects at the later years i.e "2" presents 2 years later.

Table 3.5 summarises the overall FRR at 0.01 FAR for age intervals between $0-2$ years.

Table 3.5 Comparative FRR at (0-2] years age intervals

		FRR			
	Dist.Metric / Projection	EUC	CTB	COS	MAH-COS
0.01 FAR	*PCA*	0.38	0.54	0.36	0.23
	LDA	0.24	0.26	0.25	0.14
	KPCA	0.47	0.33	0.31	0.26
	KFA	0.24	0.28	0.22	0.20

From the observation for time intervals $(0, 2]$, it can be seen that LDA algorithm has a constant False Reject Rate at 0.01 FAR with an average of 0.22. the second best performer is the KFA algorithm with an average of 0.24. Since this time intervals is fairly close to the initial enrollment age, therefore, the FRR should remain relatively low to show the effect of face ageing.

With the combination of projection algorithm and distance metrics, LDA-Mahalanobis Cosine is the best performer of 0.14. The second performer is also from the Mahalanobis Cosine with KFA with FRR of 0.20.

On another note, a further experiment for FRR at 0.005 FAR was implemented and it can be observed that a steady decrease of EER at an average of 0.02 to 0.06 was noted depending on the selected combination. This can be concluded that as the FAR increases the EER also increase in overall performances. This applied for time interval between 0-3 and 0-4 years period as well.

Table 3.6 Comparative FRR at (0-3] years age intervals

		FRR			
	Dist.Metric / Projection	EUC	CTB	COS	MAH-COS
0.01 FAR	*PCA*	0.55	0.61	0.38	0.39
	LDA	0.30	0.36	0.38	0.31
	KPCA	0.49	0.40	0.34	0.30
	KFA	0.39	0.47	0.37	0.24

From Table 3.6, it can be readily observed that as the time intervals increase from 2 to 3 years, the FRR increases as well. If we look at PCA-Mahalanobis Cosine combination, the FRR increases from 0.23 to 0.39 whereas PCA-Cosine combination has a little increment of 0.02. It can also be observed that LDA algorithm showed a constant increases in all of the combination from 0.24 to 0.30 in Euclidean, 0.26 to 0.36 in CityBlock , 0.25 to 0.38 in Cosine and 0.14 to 0.31 in Mahalanobis Cosine respectively.

This is evident that LDA-Mahalanobis Cosine and KFA-Mahalanobis Cosine are still the best performers after the 3 years intervals with the lowest FRR of 0.31 and 0.24 compared to Euclidean distance metric of 0.30 and 0.39 FRR.

Table 3.7 Comparative FRR at (0-4] years age intervals

		FRR			
	Dist.Metric / Algorithm	EUC	CTB	COS	MAH-COS
0.01 FAR	*PCA*	0.68	0.76	0.53	0.46
	LDA	0.41	0.51	0.42	0.48
	KPCA	0.58	0.75	0.59	0.62
	KFA	0.44	0.53	0.64	0.34

Table 3.7 shows the FRR values after 4 years intervals. It can be seen that a significant changes in FRR at 0.01 FAR occurred. From the PCA-Mahalanobis Cosine combination, FRR started off with 0.23 (2 years intervals) to 0.48 (4 years intervals). There is a differences of 0.25 in FRR. The best performer for this interval is KFA-Mahalanobis Cosine with a FRR of 0.34.

For the LDA-Mahalanobis Cosine combination, there is a difference of 0.20 with the initial FRR of 0.14 at 2 years intervals to 0.34 at 4 years intervals. In conclusion, as the time intervals increases, it can be evidently observed that FRR also increases.

3.9.2 Effect of ageing for age sub-groups

In this section, we will explore further on the longitudinal ageing based on "young" and "mature" age group using the same subset of 100 subjects. The "young" age group consisted of 40 subjects between $\leq 19-25$ years old whereas the remaining 60 subjects between $26- \geq 50$ years old are allocated in the "mature" age group.

Since LDA algorithm was one of the best performer during the preliminary experiments, this algorithm has been further explored in this section to see whether by separating the whole papulation into smaller groups will improve the performance.

Table 3.8 Comparative FRR for "Young" $(17-25)$ **age group**

	Age Intervals	Dist.Metric / Algorithm	EUC	CTB	COS	MAH-COS
0.01	**(0,2]**	*LDA*	0.19	0.16	0.27	0.29
FAR	**(0,4]**		0.68	0.70	0.60	0.72

From Table 3.8, the "young" group showed a huge difference as the time intervals increases from 2 to 4 years . It can be readily observed that in all the distance metrics, there are an increment of average 0.50 in the FRR. It can be understand that younger age group change more drastically in their features as they are still growing in their early life.

Table 3.9 Comparative FRR for "Mature" (26 – 65) age group

	Age Intervals	Dist.Metric / Algorithm	EUC	CTB	COS	MAH-COS
0.01 FAR	**(0,2]**	*LDA*	0.10	0.13	0.26	0.24
	(0,4]		0.32	0.37	0.35	0.41

It is interesting to observed that in Table 3.9, the "mature" age group perform in a more steady rate. There is a constant increase for all combination such as 0.10 to 0.32 in the Euclidean, 0.13 to 0.37 in the CityBlock, for Cosine its 0.26 to 0.35 and Mahalanobis-Cosine from 0.24 to 0.40. However this drop is not as severe as observed for the "young" age group.

3.10 Conclusion

This chapter presents an experimental methodology for the evaluation of the proposed approach. It covers the experimental system, feature extraction methods, projection technique, distance metrics and the databases that were used for evaluation of face ageing recognition algorithms. After the preliminary experiment, it can be concluded that "mature" age group performed better as they have a mature facial features which doesn't changes frequently compared to the "young" age group.

In the next chapter lateral ageing based on age bands will be presented.

CHAPTER 4

Face recognition across ages

4.1 Introduction

This chapter investigates the performance variation of facial recognition systems due to the influence of age. A comparative analysis of verification performances is conducted for several projection techniques combined with different distance metrics. Experimental results are based on a subset of the MORPH-II database and show that age has a significant impact on the performance of the face recognition systems investigated.

This chapter is organized as follows: The motivation of lateral ageing is briefly explained in Section 4.2. Following the preliminary experiment in previous chapter, further evaluations of the algorithms with projection techniques are included in Section 4.3.1. Section 4.3.3 presents the summary results from the experiment while Section 4.3.4 presents the findings using detection error tradeoff curve. Section 4.3.7 presents the multi-classifier or various parameter in the proposed face recognition system. Finally some concluding remarks of the chapter is provided in Section 4.4.

4.2 Motivation

The idea behind this investigation is that the biometric system has uneven difficulty in recognising people from different ages. Some algorithms may perform better for certain age groups. It is suggested that a carefully optimised system can reduce the error rates in terms of the performance. For example it has been shown that features extracted from the face by Ye et al. [165] as biometric data exhibit particular differences as a function of age.

As reported by Ramanathan et al. [109], during the formative years of a person, the variations in the shape of the face are more prominent, while in the later stage of life texture variations such as skin wrinkles and the change in pigmentation are more visible. Hence, a face recognition system which takes into account age progression needs to be able to incorporate both types of ageing factors.

In iris biometric, Erbilek et al. [166] explore the effects of different age band assignments in order to guide and enhance the management of age-related data and more objectively determine optimal age bands for system development. The authors proposed an approach to age prediction from iris images by using a combination of a small number of very simple (easily and efficiently computable) geometric features (ignoring texture-based information which is less likely to carry significant age-related information) and a more versatile and intelligent classifier structure which can achieve accuracies to 75% with three age groups (corresponding to "younger" (< 25),"middle-age" ($25 - 60$), and"older" (> 60) categories).

Ricanek et al. [15] used age groups of "18-29" , "30-39", "40-49", "$> 50''$ and also in a separate study Singh et al. [167] used age groups of "1-18" , "19-40", "> 41", while Ho et al. [168] has shown that there is no statistically significant evidence to suggest that faces of younger people are harder to identify than those of older people when using age grouping boundaries of "20", "30", "40", "50" and "60". However a study conducted by Lui et al. [169] concluded that older people are easier to recognize than younger ones. It is crucial to note that the result of obtaining lower error in older populations should be regard as degraded the recognition performance for younger population.

4.3 Experimental setup

The investigation includes various face verification systems and variations in their performance when in smaller age groups. Four different feature selection criteria and four distance metrics were studied for the matching algorithm. More details about the feature selection and distance metrics were previously discussed in Section 3.5 and Section 3.6 respectively.

100 subjects were selected from the MORPH-II database for the preliminary experiment. The subjects were then splited into five smaller age groups ("$\leq$ 19", "20-29", "30-39", "40-49", "$\geq$ 50"). In order to produce a more consistent results, 20 subjects were considered in each age band. The age range used for this study is between "17-65"(inclusive) years old. Gabor filters are used to extract features from the face images as explained in Section 3.4.

For comparative performance analysis, the verification scenario has been implemented in this study. In a biometric verification scenario, FAR denotes the portion of the impostors accepted as genuine whereas FRR denotes the proportion of genuine users rejected as impostor. The FAR and FRR depend on the operational parameters. The phenomenon when FAR equals FRR, these errors are termed as EER. All these will be reported for the comparative analysis here.

4.3.1 Experimental results

In this study, the image input into the face recognition will initially split into 5 age groupings. Each age groups will have 20 subjects with equal gender of male and female. As previously reported in Chapter 3, it can be seen that as the age progress the performance of the biometric system degraded. In this second phrase of the experiment, the population are split into 5 age bands ($\leq$ 19, 20-29, 30-39, 40-49, $\geq$ 50) depending on their first enrollment age. 20 subjects were randomly picked for each group ensuring that there are equal gender distributions in each group. There are 6 images from each person. Of these 6 images, 3 were used for enrollment and the rest for verification.

4.3.2 Effect of age using Gabor features with various projection algorithms and different distance metrics

In the following experiments, Gabor features were implemented with different types of scenario namely:

- Projection algorithm: PCA, LDA, KPCA and KFA.

- Distance Metric: Euclidean, CityBlock, Cosine Similarity and Mahalanobis-Cosine distance.

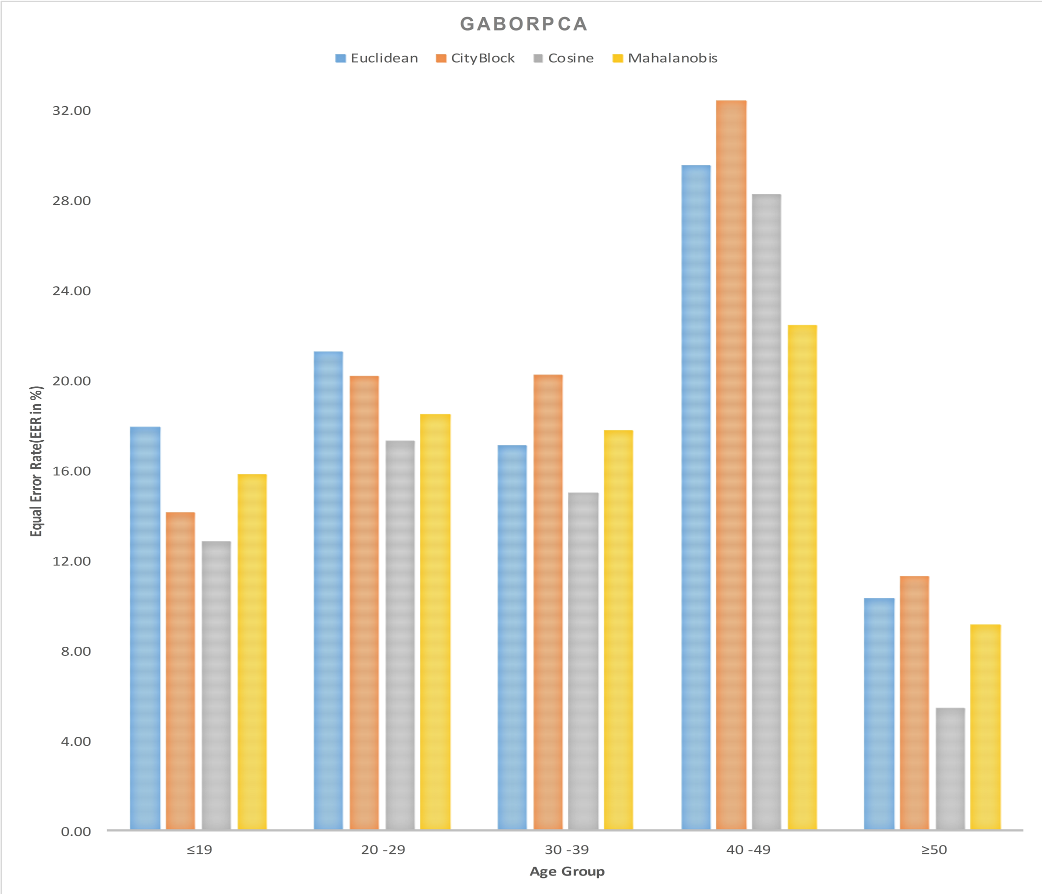

Figure 4.1 Effect of GaborPCA based on distance metric with age grouping

In this experiment, we investigate the PCA algorithm with various distance metric combinations for the chosen age groups was considered. Figure 4.1 presents the corresponding EER values.

For all the age groups except "40-49" years, it can be observed that GaborPCA-Cosine combination has the lowest EER compared to the other 3 distance metric combination. The "40-49" group was the most difficult and lower performer. However, the Mahalanobis-Cosine distance is the best performer with EER of 22.5%. Among all the combinations for GaborPCA, the best accuracy achieved was 5.49% for the " $\geq 50''$ when Cosine distance combination were applied.

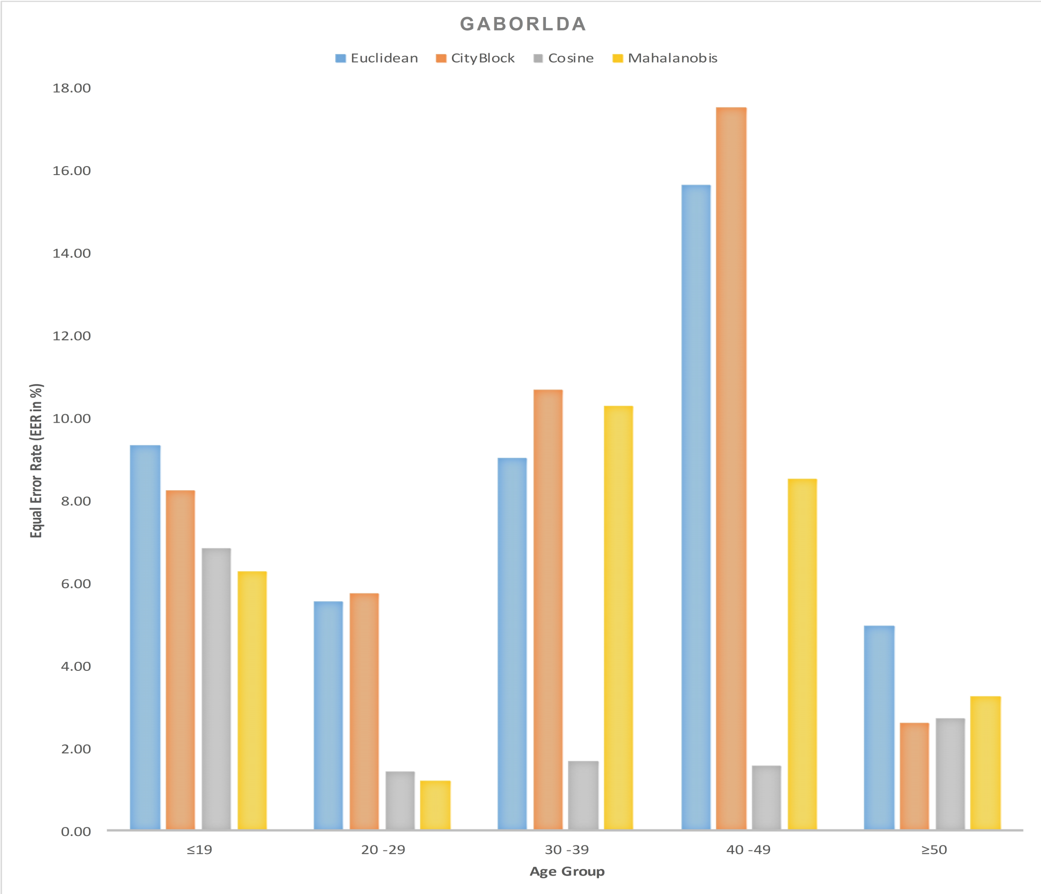

Figure 4.2 Effect of GaborLDA based on distance metric with age grouping

Figure 4.2 shows the EER for LDA with distance metric combination. From the observation, for age group "$\leq$ 19" and "20 - 29", the Mahanalobis-Cosine distance produced the best performance with EER of 6.3% and 1.2% respectively. For the remaining 3 age groups ("30 -39", "40-49" and "$\geq$ 50"), the Cosine distance has lower EER. It can also be seen in Figure 4.2 that Euclidean and CityBlock distance has a disproportionately higher EER when compared to Cosine and Mahalanobis distance combinations.

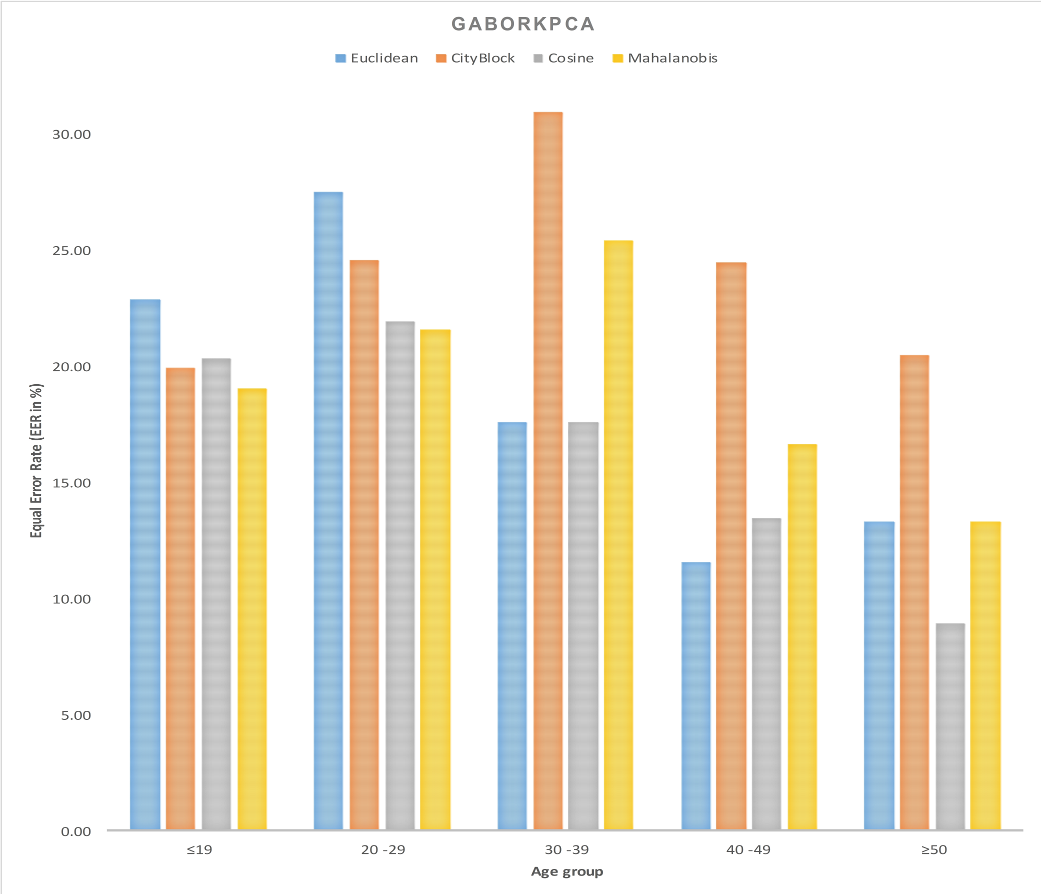

Figure 4.3 Effect of GaborKPCA based on distance metric with age grouping

It can readily be observed that KPCA did not perform as good as the other 3 projection algorithms as shown in Figure 4.3. But nevertheless, it can be noticed that as the age progresses, the KPCA performs relatively better. Age group "40 -49" has an EER of 11.6% in Euclidean distance combination when compared to age group of " $\leq 19''$ with an EER of 22.9%. For the age group of " $\geq 50''$, the Cosine distance combination has the lowest overall EER which is 8.9%.

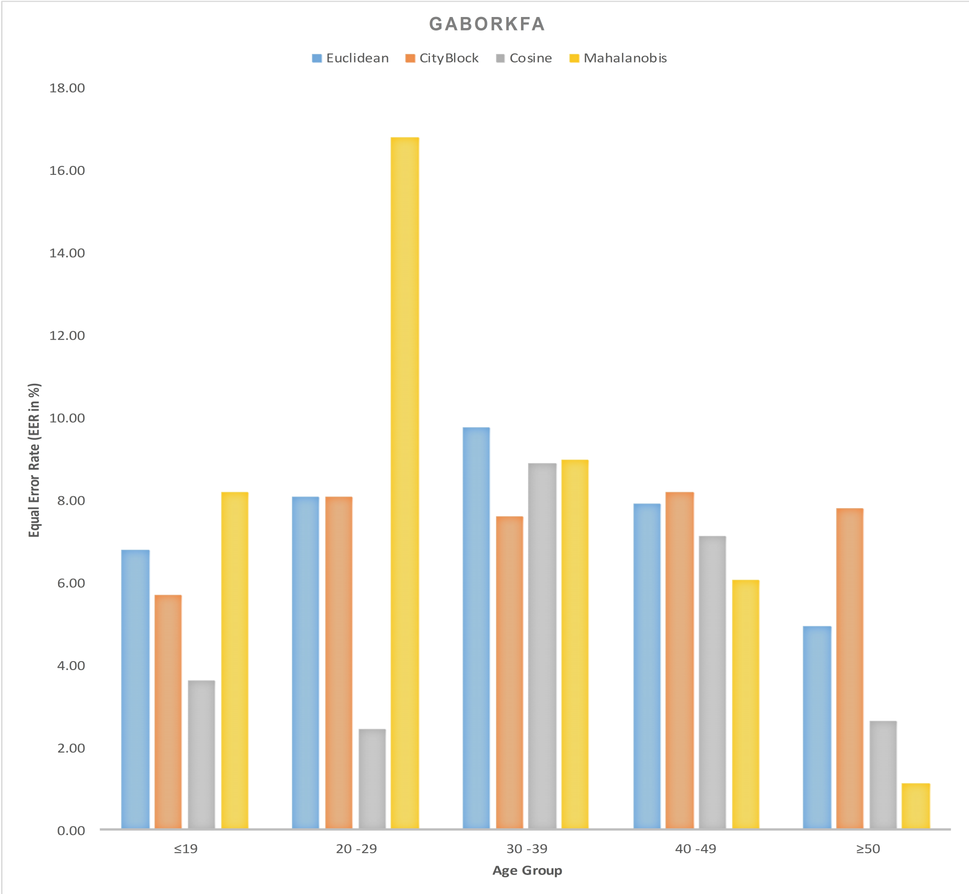

Figure 4.4 Effect of GAborKFA based on distance metric with age grouping

The last projection algorithm considered was the second best performer after LDA. Figure 4.4 presents the GaborKFA algorithm with different distance combinations.

In these observations, all the 4 mentioned distance metrics performed steadily for all of the tests. For KFA, the 2 age groups that had the best performance were the youngest group "20-29" with an EER of 2.5% on Cosine distance combination and the most mature group "$\geq$ 50" with an EER of 1.1% on Mahalanobis-Cosine distance combination.

Between age group of “40-49” and “$\geq$ 50” , it can be clearly observed that mature population obtained a lower EER compared to age group below 40 years old. This can be evidently mentioned that mature people have a steady facial changes throughout the 5 years when comparable to younger population (i.e gender differences when growing up). It can also be noted that there are a few types of ethnicity in this dataset which can be linked to their different lifestyle as previously reported in [77].

4.3.3 Summary for the projection algorithms versus distance metrics

Table 4.1 shows the EER for all possible combinations of system parameters. It can readily be observed that, even for the same combination of selection algorithm and distance metrics, the EER values for each age group varied quite drastically. As reported in the previous chapter, the LDA-Cosine Distance combination produced the lowest EER of 4% when there was no age grouping; but with the age grouping incorporated, the EER values for the groups were significantly lower (the lowest being 1.4% for the “20-29” age band) except for the “$\leq 19''$ group where the EER increased to 6.3%. Even for the KPC-Mahalanobis/Cosine combination which originally produced a poor 49.5% EER, after age grouping the EER values ranged between 13.3% and 25.5%.

In the “20-29” age grouping, GaborKPCA and GaborKFA performed well, on all recognition task. Mahalanobis-Cosine is the best performing distance metric with all the algorithms. It is interesting to acknowledge that both non-linear subspace projection techniques are the best performers in this group.

For the “30-39” age grouping, Cosine and Mahalanobis-Cosine distance metric performed well on all recognition tasks with a low error rate of 1.7% and 9.0% respectively. GaborKFA can be considered the best performing algorithms in “30-39” age grouping.

In the “40-49” age grouping, Euclidean and Cosine distance metrics are generally the best performers as they each achieved lowest EER in 3 out of the 4 projection algorithms. They performed relatively well on both GaborLDA and GaborKPCA algorithms. It is worth to elaborate this is the only age group that works well with Euclidean distance metric as this distance perform satisfactory average on all the experiments.

For the age group of "$\geq$ 50" years old, Cosine and Mahalanobis-Cosine distance metrics are the best performers with the lowest EER of 2.6% and 1.1% respectively. Cosine and Mahalanobis-Cosine are the best performers distance metrics for all age grouping. For Euclidean distance metric, this combination was considered the best for "40-49" age group.

It is also noticed that no particular combination produced the best EER for all the age bands (the lowest EER for any age band is shown in bold in Table 4.1 suggesting that, for optimum performance, different schemes should be used for different age groups. It can also be noticed that, of all the age groups, the "$\leq$ 19" age group showed the highest EER suggesting that this age group is experiencing most variation in their facial appearances.

Table 4.1 Comparative EER for different age groups

Projection Algorithm	**Distance Metric**	**Age Group**				
		≤ 19	20-29	30-39	40-49	≥ 50
Gabor PCA	*Euclidean*	0.18	0.21	0.17	0.29	0.10
	CityBlock	0.14	0.20	0.20	0.32	0.11
	Cosine	0.13	0.17	0.15	0.28	0.05
	Mahalanobis- Cosine	0.16	0.19	0.18	0.23	0.09
Gabor LDA	*Euclidean*	0.09	0.05	0.09	0.16	0.05
	CityBlock	0.08	0.06	0.11	0.18	0.03
	Cosine	0.07	0.01	**0.02**	**0.02**	0.03
	Mahalanobis- Cosine	0.06	**0.01**	0.10	0.09	0.03
Gabor KPCA	*Euclidean*	0.23	0.28	0.18	0.12	0.13
	CityBlock	0.20	0.25	0.31	0.25	0.21
	Cosine	0.20	0.22	0.18	0.14	0.09
	Mahalanobis- Cosine	0.19	0.22	0.26	0.17	0.13
Gabor KFA	*Euclidean*	0.07	0.08	0.10	0.08	0.05
	CityBlock	0.06	0.08	0.08	0.08	0.08
	Cosine	**0.04**	0.03	0.09	0.07	0.03
	Mahalanobis- Cosine	0.08	0.17	0.09	0.06	**0.01**

4.3.4 Detection Error Tradeoff (DET)

Equal Error Rates (EER) presents a snapshot of the system performance for a particular operating threshold. To show the general trend in the variations of FAR-FRR over a range of thresholds, the DET-curve is an ideal tool.

In this section, we shall look at the False Accept Rate (FAR) - False Reject Rate (FRR) over a range of thresholds. Figure 4.5 presents an example of DET curve at different thresholds extracted from the KPCA-Cosine combination.

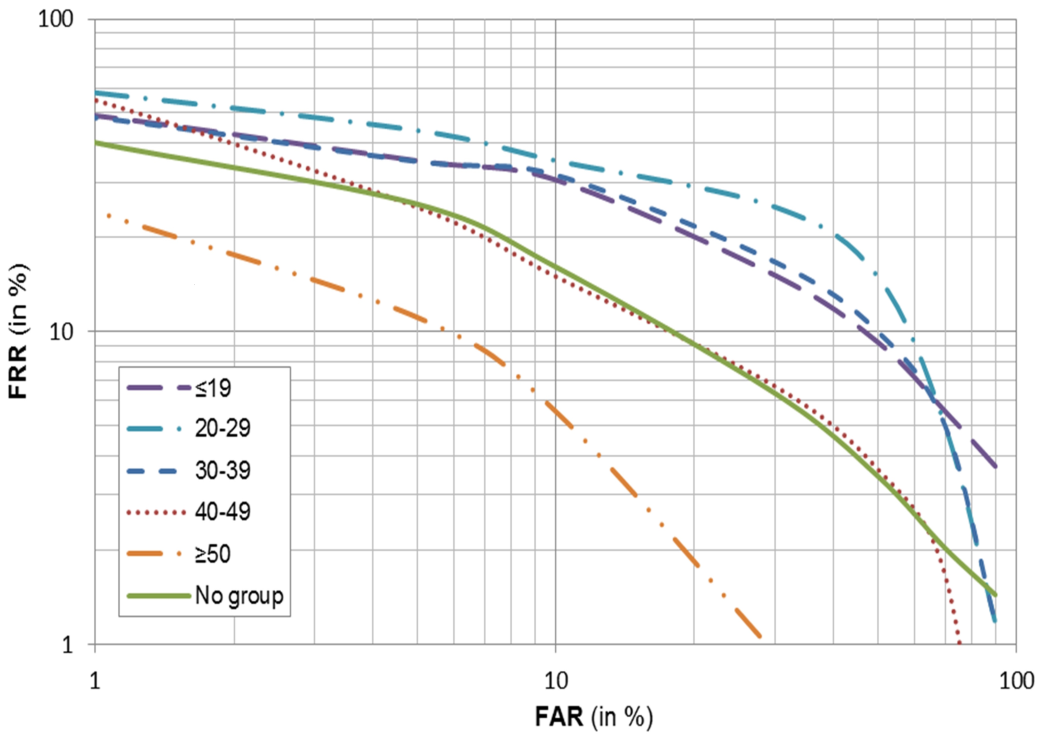

Figure 4.5 DET curve for the KPCA-Cosine combination

From Figure 4.5, the "$\geq$ 50" age group consistently performs better for all operating thresholds. The "40-49" age group performance was comparable to that of "No Grouping". The performances for the remaining age groups are relatively less accurate as compare to the "no grouping" curve.

4.3.5 Effect of age groups based on Rank-1 recognition rates

In this section, the experiments will further explored the effect of face ageing. The first experiments will look into Rank-1 recognition rate for Gabor with projection algorithms combination.

Table 4.2 Comparative analysis for projection algorithm Rank-1 recognition rates

Feature Extraction	Age Group				
Types	$\leq$ 19	20 -29	30 -39	40 -49	$\geq$ 50
PCA	0.70	0.90	0.56	0.81	0.90
LDA	0.90	**1.00**	0.71	0.81	**1.00**
KPCA	0.90	0.99	0.80	**0.90**	1.00
KFA	**0.91**	0.80	**0.90**	0.81	0.90

The subjects for each age group are trained in a one-to-many (1:M) scenario. Each subject consisted of 6 images, where 1 image was used for enrollment and the remaining were used for identification). In Table 4.2, the recognition rates vary among the age groups. GaborKFA performed best in the age group of "$\leq$ 19" and "30-39" with the recognition rates of 90.52% and 90.20% respectively. GaborLDA has a full 100% rank-1 recognition rates for age group "20-29" and "$\geq$ 50". All of the best performance are highlighted in bold.

The rank-1 recognition rates for all of the age groups recorded high values, it should be noted that age group "40-49" is the only group that has the lowest recognition rate in both GaborKFA and GaborLDA algorithms. GaborKFA combinations has a consistent high values for the 5 age groups which ranged between 80.34% to 90.52% as shown on Table 4.2. In terms of perfect score, GaborLDA combinations has 100% for both age groups "20-29" and "$\geq$ 50" respectively.

It can be observed when ageing is incorporated into the system, for the younger population of age group "$\leq$ 19" and "20-29", LDA has a relatively higher recognition rates compared to PCA in the linear techniques using Gabor method. As LDA was also the best performer for the oldest age group "$\geq$ 50", this feature extraction was further explored in Section 4.3.7.

4.3.6 Effect on number of enrollment images on projection performance

In this section, the number of enrollment images used for each experiment varies to show the effect on the EER for different age group. In these results, "n" is the number of images used for enrollment. For example when $n = 2$, out of these 6 images, 2 were used for enrollment and the rest for verification. EER values for GaborPCA classifier are shown on the Table 4.3 with different distance metric combinations. The EER values can be seen as a constant progress when there is more images used in the enrollment.

In the GaborPCA-Cosine combination, when number of enrollment images increase from 2 to 4 for the age group "$\leq$ 19", the EER value decreases from 19.82% to 14.25%. In the age group of "20-29", the EER value decreases from 19.55% to 17.58%. It can be concluded that as the number of enrollment images increases more than 3 images per testing, the overall EER value will perform a lot better than with smaller number of enrollment images.

While observing at different number of images per enrollment, i.e. when n=2, the lowest EER value for 3 particular age groups (i.e. "$\leq$ 19" , "20-29" and "$\geq$ 50" with all three of them belonged to Mahalanobis-Cosine distance combination with 14.23% ("$\leq$ 19") , 18.67% ("20-29") and 8.42%("$\geq$ 50") respectively. For the age group of "30-39", the Euclidean distance combination produce the lowest EER of 19.35%. Meanwhile, for the age band of $40 - 49$, the lowest EER are recorded for the Cosine distance combination with EER value is 27.83%.

When the number of enrollment images increases to n=3, all of the the age groups perform the best in the Mahalanobis-Cosine distance combination except for the "40-49" age group (with Cosine distance combination). The performance for this 3 images per enrollment is comparatively lower than 2 images per enrollment with the difference in EER of 0.89% ("$\leq$ 19"), 1.80%("20-29"), 3.93%("30-39") and 2.66%("$\geq$ 50"). The age group of "40-49" has a difference EER values of 2.07% when compared to a lower number of images used in enrollment.

In the last setup when n=4 where there are more than half of the total images per subject, Mahalanobis-Cosine distance combination produced the lowest EER value for all the age groups among the rest of the possible parameters. It can be observed there is a significant decrease in the EER when compared to small dataset used for enrollment. The EER for the following age group are as follows:

"$\leq$ 19" age group with an EER value of 11.18% , "20 -29" age group with an EER value of 14.17%
"30 -39" age group with an EER value of 13.78%, "40 - 49" age group with an EER value of 12.63% and "$\geq$ 50" age group with an EER value of 5.38%.

Table 4.3 Comparative EER with number of enrolled images for GaborPCA

GaborPCA with Distance Metric	no. of enrolled image	Age group				
		≤ 19	20 - 29	30 - 39	40 - 49	≥ 50
Euclidean	n = 2	0.21	0.23	0.19	0.36	0.10
	n = 3	0.18	0.21	0.17	0.30	0.10
	n = 4	0.16	0.18	0.17	0.18	0.08
CityBlock	n = 2	0.23	0.24	0.32	0.34	0.13
	n = 3	0.14	0.20	0.20	0.32	0.11
	n = 4	0.13	0.19	0.18	0.17	0.11
Mahalanobis -Cosine	n = 2	0.14	0.19	0.28	0.31	0.08
	n = 3	0.13	0.17	0.15	0.28	0.06
	n = 4	**0.11**	**0.14**	**0.14**	**0.13**	**0.05**
Cosine	n = 2	0.20	0.20	0.22	0.28	0.12
	n = 3	0.16	0.19	0.18	0.23	0.09
	n = 4	0.14	0.18	0.17	0.22	0.08

Table 4.4 Comparative EER with number of enrollment image for GaborKPCA

GaborKPCA with Distance Metric	no. of enrolled image	Age group				
		≤ 19	20 - 29	30 - 39	40 - 49	≥ 50
Euclidean	n = 2	0.24	0.28	0.19	0.17	0.13
	n = 3	0.23	0.28	0.18	0.12	0.13
	n = 4	0.21	0.16	0.17	0.11	0.09
CityBlock	n = 2	0.21	0.26	0.33	0.26	0.23
	n = 3	0.20	0.25	0.31	0.25	0.21
	n = 4	**0.10**	0.22	0.21	0.19	0.15
Mahalanobis -Cosine	n = 2	0.22	0.24	0.39	0.18	0.15
	n = 3	0.19	0.22	0.26	0.17	0.13
	n = 4	0.16	**0.13**	0.19	0.13	0.08
Cosine	n = 2	0.25	0.23	0.20	0.16	0.19
	n = 3	0.20	0.22	0.18	0.14	0.09
	n = 4	0.20	0.13	**0.15**	**0.10**	**0.06**

GaborKPCA classifier is shown on Table 4.4 with the distance metrics combination. The EER values can be seen as a constant progress when there is more images used in the enrollment. While observing at different number of images per enrollment, i.e. when $n = 2$, the lowest EER for age group "20-29" and "40-49" are in Cosine distance combination with 23.38% and 15.83% respectively. For age group of "30-39" and "$\geq$ 50" , the Euclidean distance combination produce the lowest EER value of 19.21% and 13.02% respectively. Meanwhile, for age group of "$\leq$ 19", the lowest EER are recorded in the CityBlock distance combination with EER value of 21.56%.

When the number of enrollment images increases to $n = 3$, all the age groups performed well in the Euclidean and Cosine distance combinations except for the "$\leq$ 19" age group (with Mahalanobis-Cosine distance combination). The performance for this enrollment images is comparatively lower than the previous enrollment (where $n = 2$) with the difference in EER of 1.63% for "20 -29" age group, 0.84% for "30 -39" age group, 3.68% for "40-49" age group and 3.77% for age group "$\geq$ 50". The "$\leq$ 19" age group has a difference EER value of 2.81% when compared to the lower number of images used in enrollment.

In the last setup when $n = 4$ where there is more than half of the total images in the testing, Cosine distance combination produced the lowest EER values for all the age groups except for age group "$\leq$ 19" (with CityBlock distance combination). It can be observed there is a significant decrease in the EER value when compared to smaller dataset used for enrollment. The EER for all of the age group are as follows:

- "$\leq$ 19" age group with an EER value of 10.27%
- "20-29" age group with an EER value of 13.38%
- "30-39" age group with an EER value of 14.66%
- "40-49" age group with an EER value of 10.25% and
- "$\geq$ 50" age group with an EER value of 5.74%.

Table 4.5 Comparative EER with number of enrollment image for GaborKFA

GaborKFA with Distance Metric	no. of enrolled image	Age group ≤ 19	20 - 29	30 - 39	40 - 49	≥ 50
Euclidean	n = 2	0.17	0.09	0.11	0.08	0.10
	n = 3	0.07	0.08	0.10	0.08	0.05
	n = 4	0.16	0.06	**0.05**	0.07	0.02
CityBlock	n = 2	0.17	0.09	0.09	0.09	0.12
	n = 3	0.06	0.08	0.08	0.08	0.08
	n = 4	0.05	0.06	0.07	0.05	0.02
Mahalanobis -Cosine	n = 2	0.09	0.25	0.10	0.07	0.03
	n = 3	0.08	0.17	0.09	0.06	0.02
	n = 4	0.05	0.16	0.08	0.06	**0.01**
Cosine	n = 2	0.09	0.04	0.11	0.08	0.04
	n = 3	0.04	0.03	0.09	0.07	0.03
	n = 4	**0.04**	**0.02**	0.08	**0.02**	0.02

Table 4.5 represents GaborKFA classifier with various distance metric combination. When $n = 2$, the lowest EER value for "≤ 19", "40-49" and "≥ 50" age groups are in Mahalanobis-Cosine distance combination with 9.42%, 7.36% and 2.87% respectively. For the age group of "20-29", the Cosine distance combination produced a low EER value of 4.46%. Meanwhile, for the age group of "30-39", the lowest EER value was noted in the CityBlock distance combination with 9.42%.

When the number of enrolled images increases to $n = 3$, all the age groups performed the best for both Mahalanobis-Cosine and Cosine distance combinations except the "30-39" age group (with CityBlock distance combination). The performance for $n = 3$ enrolled images are comparable lower to $n = 3$ with the differences of 5.04% ("≤ 19" age group), 1.14% for ("20-29" age group) , 1.24% for("40-49" age group) and 1.72% for age group "≥ 50". The "30-39" age group has a difference EER of 1.21% when compared to the lower number of images used in enrollment.

In the final setup when $n = 4$ where there is more than half of the total images per subject, both Mahalanobis-Cosine and Cosine distance combinations produced the lowest EER for all age groups except for "30-39" (Euclidean distance combination)age group. It can be observed there is a significant decrease in the EER when compared to small dataset used for enrollment. The difference of EER between $n = 2$ and $n = 4$ for all of the age groups are as follows:

- "≤ 19" age group with an EER value of 5.05%
- "20-29" age group with an EER value of 2.60%
- "30-39" age group with an EER value of 4.13%
- "40-49" age group with an EER value of 5.03% and
- "≥ 50" age group with an EER value of 1.76%.

Table 4.6 Comparative EER with number of enrollment image for GaborLDA

GaborLDA with Distance Metric	no. of enrolled image	Age group				
		≤ 19	20 - 29	30 - 39	40 - 49	≥ 50
Euclidean	n = 3	0.09	0.06	0.09	0.16	0.05
	n = 4	0.08	0.04	0.02	0.08	0.02
CityBlock	n = 3	0.08	0.06	0.11	0.18	0.03
	n = 4	0.07	0.02	0.04	0.05	0.02
Mahalanobis - Cosine	n = 3	0.06	0.02	0.10	0.09	0.03
	n = 4	**0.05**	**0.01**	0.08	0.02	**0.01**
Cosine	n = 3	0.07	0.02	0.02	0.02	0.03
	n = 4	0.06	0.01	**0.01**	**0.01**	0.01

Table 4.6 represents the GaborLDA classifier with all the possible parameters on the distance metrics combinations. In this particular experiment, the number of subjects has to start with $n = 3$, as GaborLDA classifier only works with a minimum of 3 images in the enrollment stage.

LDA is seen as one of the best EER performer compared to some of the other projection algorithms. There are only 2 distance metrics that performed the best using GaborLDA classifier namely Mahalanobis-Cosine and Cosine distance combinations. When the basis vector is $n = 3$, the age group for "$\leq$ 19" produce an EER of 6.37% but as the number of vector increases to $n = 4$, the EER decreases to 4.75%. For the remaining age band, the best performer is Cosine distance combination, the differences for EER between each basis vector are as follows: "$20 - 29$" , "$30 - 39$" , "$40 - 49$" and "$\geq$ 50" with EER values of 0.67%, 1.27%, 1.05% and 2.17% respectively.

For the linear techniques, both PCA and LDA performed relatively well with Mahalanobis-Cosine distance combination regardless of the age group. Overall, age group "$\leq$ 19" produced the lowest EER in KFA - Cosine distance combination. This can be seen that non-linear technique will work better in the younger age group when facial changes more frequent than mature facial expression whereas linear technique works best with selected distance metric. It is also worth mentioning that both non-linear techniques i.e KPCA and KFA performed better with Cosine distance combination for the "young" age groups i.e "$\leq$ 19" and ""$20 - 29$" when compared to other distance metrics.

In conclusion, there are 2 best performers when age groups are taken into account. For the age group of "$\leq$ 19" Gabor-KFA demonstrated the lowest EER of 3.28% value when $n = 4$ while the remaining 4 age groups "20-29", "30-39" , "40-49" and "$\geq$ 50" had the lowest EER values in the Gabor-LDA classification.

4.3.7 Multi-Classifier Fusion

It was proven so far that different classification schemes performed better for different age groups with different settings. The following experiment shall explore whether a multi-classifier approach can produce even better accuracy rates in face ageing.

In the multi-classifier fusion scheme, features extracted from each image are used to obtain different projection algorithms and a matching. The scheme is illustrated in

Figure 4.6. In this multi-projection algorithm system, there will be various experiments to show the effect of face ageing when different types of projection algorithms are implemented in each experiments. There are two types of score fusion that had been investigated in the following experiments i.e. sum and product rule.

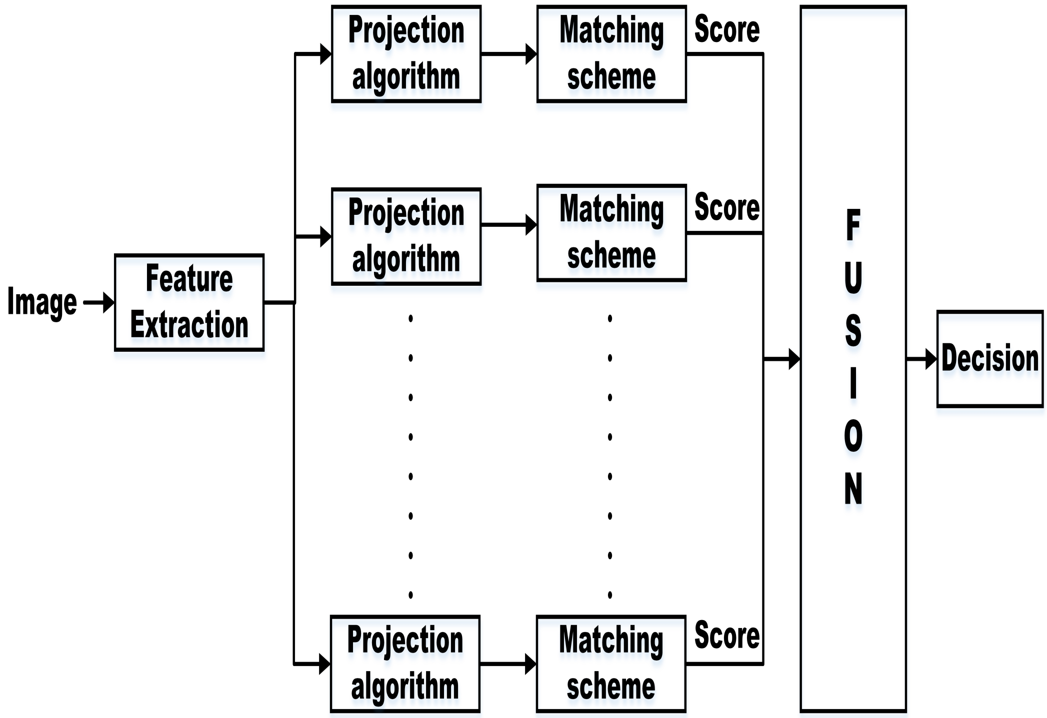

Figure 4.6 Feature fusion performance from different projection algorithms

4.3.8 Effect of age using multi-classifier fusion with various projection algorithms and different distance metrics

In the first experiment, Table 4.7 shows the EER values for 3 distance metric combinations versus different types of projection algorithms. The lowest EER obtained are highlighted in bold. From the observation, KFA is the best performer for product rule. In this Table, it can be observed that each age groups produce a noticeable lower EER values when compared to single classifier as shown in Table 4.1 except for " $\leq 19''$ which have an EER of 2.6%.

Table 4.7 Comparative EER with 3 distance metric combinations

Distance Metric	Projection Algorithm	Fusion Rule	Age Group				
			≤ 19	20 -29	30 -39	40 -49	≥ 50
Cosine Mahalanobis -Cosine Euclidean	PCA	Sum	0.182	0.153	0.157	0.158	0.072
		Product	0.143	0.161	0.149	0.148	0.079
	LDA	Sum	0.095	0.040	0.046	0.050	0.048
		Product	0.069	0.027	0.027	0.025	0.022
	KPCA	Sum	0.261	0.232	0.213	0.173	0.067
		Product	0.239	0.196	0.187	0.158	0.071
	KFA	Sum	0.067	0.041	0.040	0.018	0.012
		Product	**0.062**	**0.009**	**0.009**	**0.011**	**0.008**

For the following experiment in Table 4.8, it can be observed that KFA algorithm is the best performer. It can be noted that both sum and product rule. The highest difference of EER value is "30-39" age group which is lowered by 1.1% when compared to Table 4.1. The remaining age groups (i.e. "20-29", "40-49" and "$\geq$ 50") presents a lower EER values as well when compared to Table 4.1. The lowest EER value for age group "$\geq 50''$ with 0.3% lower when comparing with Table 4.1. This Table also shows the lowest overall EER value for age group "40-49" with 0.011 when compared to the remaining experiments.

Table 4.8 Comparative EER with 2 distance metric combinations

Distance Metric	Projection Algorithm	Fusion Rule	Age Group				
			≤ 19	20 -29	30 -39	40 -49	≥ 50
Cosine Mahalanobis -Cosine	PCA	Sum	0.149	0.153	0.154	0.151	0.079
		Product	0.148	0.151	0.153	0.150	0.074
	LDA	Sum	0.059	0.041	0.034	0.030	0.021
		Product	0.056	0.033	0.023	0.027	0.020
	KPCA	Sum	0.197	0.199	0.198	0.169	0.068
		Product	0.195	0.196	0.197	0.171	0.067
	KFA	Sum	**0.055**	0.007	0.008	0.017	0.006
		Product	0.057	**0.007**	**0.007**	**0.016**	**0.005**

In the next experiment, Table 4.9 shows the EER values for all distance metric combinations with 3 different types of projection algorithms. From the observation, CityBlock and Euclidean distance combinations are the best performer for Product rule. In the Table 4.9 below , it can be observed that each age group illustrated a noticeable lower EER values when compared to no-fusion classifier as shown in previous section, Table 4.1.

When 2 distance metrics are fused together, age groups "20-29", "30-39" and "$\geq$ 50" managed to produce the lowest EER values when compared to Table 4.7, Table 4.9 and Table 4.10.

Table 4.9 Comparative EER with 3 projection algorithms

Projection Algorithm	Distance Metric	Fusion Rule	Age Group				
			≤ 19	20 -29	30 -39	40 -49	≥ 50
PCA KPCA KFA	Cosine	Sum	0.096	0.072	0.071	0.065	0.032
		Product	0.098	0.063	0.060	0.061	0.030
	Mahalanobis -Cosine	Sum	**0.081**	0.064	0.063	0.066	0.020
		Product	0.083	0.060	0.059	0.061	0.019
	CityBlock	Sum	0.253	0.252	0.252	0.251	0.125
		Product	0.103	**0.052**	**0.046**	**0.060**	0.032
	Euclidean	Sum	0.260	0.230	0.231	0.172	0.066
		Product	0.095	0.073	0.057	0.065	**0.017**

In the following experiment, all of the projection algorithms are being investigated for all possible combinations of the system parameters. Table 4.10 shows the EER score fusion for the 4 distance metrics combined with 4 projection algorithms. It can be readily observed that 4 projection algorithms did not perform as good as the other fusion classifiers compared to Figure 4.9.

Nevertheless, in the Euclidean distance combination, it can be seen that product rule produced a lower EER in all age bands compared to sum rule. The highest difference of EER value is "≥ 50" age group which is increased by 0.7% as compared to Table 4.1. The remaining age groups presents a slight increase in EER values. The remanding 4 age groups performed differently in other distance metrics combination.

Table 4.10 Comparative EER with 4 projection algorithms

Projection Algorithm	Distance Metric	Fusion Rule	Age Group				
			≤ 19	20 - 29	30 - 39	40 - 49	≥ 50
PCA LDA KPCA KFA	Cosine	Sum	0.080	0.040	0.034	0.046	0.019
		Product	0.079	0.037	0.034	**0.044**	0.017
	Mahalanobis -Cosine	Sum	0.077	0.036	0.036	0.049	0.010
		Product	**0.076**	0.031	0.031	0.045	0.009
	CityBlock	Sum	0.252	**0.252**	0.253	0.245	0.252
		Product	0.093	0.050	0.033	0.040	0.093
	Euclidean	Sum	0.259	0.227	**0.2288**	0.169	0.006
		Product	0.082	0.048	0.031	0.056	**0.004**

When compared Table 4.10 with Table 4.9, the EER values for this Table is noticeably lower in some age groups especially in "≥ 50" age group and it shows a significant improvement in the face recognition system. This is evidently proved that with 4 projection algorithms fusion classifiers, after comparing with Table 4.9 the overall EER values lower by an average of 1.05%.

In the next experiment of score fusion, Table 4.11 represents the comparative EER with 2 projection algorithms. This table illustrate the lowest EER values for age group " ≤ 19" and "≥ 50" when compared to Table 4.1, Table 4.7, Table 4.8, Table 4.9

and Table 4.10. The 3 best performers utilise with Mahalanobis-Cosine, Cosine and Euclidean distance combinations. In comparison to Table 4.1,the differences of EER for all the age groups are as follows: ≤ 19 age group with a difference EER of 0.95%, "20 -29" age group with a difference EER of 0.14%, "30 -39" age group with a difference EER of 0.86%, "40 -49" age group with a difference EER of 0.33% and ≥ 50 age group with a difference EER of 0.58%.

Table 4.11 Comparative EER with 2 projection algorithms

Projection Algorithm	Distance Metric	Fusion Rule	Age Group				
			≤ 19	20 -29	30 -39	40 -49	≥ 50
LDA KFA	Cosine	Sum	0.039	**0.008**	0.032	0.024	0.017
		Product	**0.030**	0.008	0.031	0.405	0.015
	Mahalanobis -Cosine	Sum	0.067	0.024	0.011	0.016	0.019
		Product	0.068	0.026	**0.011**	**0.016**	0.018
	CityBlock	Sum	0.093	0.044	0.038	0.037	0.043
		Product	0.091	0.038	0.020	0.029	**0.014**
	Euclidean	Sum	0.087	0.039	0.038	0.022	0.019
		Product	0.081	0.040	0.040	0.038	0.012

Table 4.12 Summary of age groups using different parameters in the system

Types of Fusion		Age Group				
		≤ 19	20 -29	30 -39	40 -49	≥ 50
No Fusion		0.036	0.012	0.017	0.016	0.011
Projection Algorithm Fusion	PCA, LDA, KPCA & KFA	0.076	0.050	0.033	0.040	0.015
	PCA, KPCA & KFA	0.081	0.050	0.046	0.060	0.017
	LDA & KFA	**0.030**	0.008	0.011	0.016	0.014
Distance Metric Fusion	Cosine, Mah-Cos & Euclidean	0.062	0.009	0.009	**0.011**	0.008
	Cosine & Mah-Cos	0.055	**0.007**	**0.007**	0.016	**0.005**

Table 4.12 presents the summary of EER for all possible combinations of system parameters. The EER results represents that with multi-classifier implementation, the system performance improve to nearly 99% to all age groups except age group "≤ 19" and "40-49". This can be concluded that younger age group i.e "≤ 19" has the lowest accuracy rate of 96.95% when compared the older age groups (i.e between age 20 -65) has a higher accuracy between 98.9% to 99.5%.

Among all the possible combinations of system parameter on the fusion rule, distance metric fusion is the best performers for all age groups except "≤ 19" age group. It was proven that in the previous experiments, both Cosine and Mahalanobis-Cosine have the lowest EER for all the age group. Hence both of these distance metrics are combined in the fusion rule to enhance the overall performance. It can be observed that when fusion rule are incorpporated into the system, the EER is further lower by an average of 0.005 for all age groups except "$\leq$ 19". For the age group of "$\leq$ 19", projection algorithm fusion (LDA and KFA) are the best performers. This can be noted that younger age group work better in projection algorithm fusion while older group perform better in the distance metric fusion combinations. It can be seen that other than (LDA and KFA) fusion setup, the rest of the similar fusion setup in the projection algorithm obtained a higher EER when compared to no fusion was introduced. As observed in the individual experiment, PCA and KPCA have a higher EER when compared to LDA and KFA. It can be further evidently proven in the fusion setup, whenever there are PCA and KPCA combinations, the EER increases dramatically when compared to LDA and KFA combinations.

Table 4.13 Comparison of performance reports (EER) for 0-3 years

			Age Group				
Method	Author	Time lapse	<18	18-29	30-39	40-49	50+
LBP	Ricanek [16]	0 - 3	0.289	0.209	0.197	0.218	0.225

In Table 4.13, Ricanek used MORPH-II database with selected subjects using EER identification method. The author does not mentioned the number of subjects in their experiment. The patch size of each image is 15×15 with a feature vector length of 4779. In this case, implementation deployed CityBlock as the matching score.

Table 4.14 Comparison of performance reports (EER)for 0-5 years

			Age Group				
Method	Author	Time lapse	< 18	18-29	30-39	40-49	50+
PCA	Ricanek [15]	0 - 5	0.344	0.42	0.452	0.80	-
Gabor	Our method	0 - 5	0.030	0.007	0.007	0.011	0.004

From the Table 4.14, Ricanek used MORPH-I database which consisted of 515 subjects. There were 420 males and 95 females, but the experiment does not mention the number of training and testing samples per subjects. Their observation shows that the PCA algorithm performed better in identifying older people with the same age span than younger ones.

Table 4.14 shows a comparison of our experimental results with the performances reported for similar age grouping published in the literature. Although the results are based on different feature extraction approach, the author indicate the relative promise of their proposed methods. The performance of our approach can be seen to compare favourably with other feature method and lends support to its applicability to minimise the effect of face ageing.Additional research is warranted to validate these findings in the next chapter.

4.4 Conclusion

The aim of this study is to investigate and analyze the relative performance of different subspace projection algorithms combined with various distance metrics for different age groups. The experiment used a facial image database containing face images acquired over a period of five years. It was observed that system performance improved significantly when age groupings were introduced and separate matching schemes were deployed for each age group. It was also observed that for this dataset the EER and FRR were highest for the younger population ("$\leq$ 19" year olds).

In the multi-classifiers experiments, it can be concluded that with 2 projection algorithms combined, both age groups "20-29" and "30-39" achieved a lower EER, while combination of distance metrics in the fusion classifier produced a reasonable lower EER in the remaining age groups.

To summarize the above discussion, GaborKFA and GaborLDA performs well in the Gabor based algorithms, whereas Cosine and Mahalanobis-Cosine distance metrics works best in the ageing task. Further work will explore how an adaptive system can be designed to harness this age dependency for maximum recognition performance. In the next chapter new algorithm called longitudinal face ageing will be presented.

CHAPTER 5

Template Ageing using Linear Transformation

5.1 Introduction

This chapter explores the effect of longitudinal ageing on the performance of face biometric systems. There are a number of real-world applications (e.g. border security) where a system must be able to reliably match individuals over a number of years. An understanding of the ageing process is becoming an increasing important issue in terms of ensuring reliability in the face of changing data scenarios. In fact, both physical ageing (the effect of the passage of time on the biometric measurements used in the identification process) and template ageing (changes occurring across the time elapsed between enrolment; when the template is created and used; when the current measurement values are matched against those stored on the template) are important in a practical context, and may have a variety of implications for practical biometric systems. For example, in the iris-based biometric studies by Erbilek et al. [166] , it has been clearly demonstrated that physical ageing effects on iris-based biometrics are primarily the result of the physiology of pupil mechanisms, with pupil dilation responsive decreasing with age, a mechanism which actually tends to improve the performance of iris biometric system.

A novel template ageing scheme by linear transformation is proposed here. The idea behind the proposed template ageing is to create a new estimated feature vector and compare with the image to be verified. This idea is extensively investigated using two recognition systems. Datas were taken from the MORPH-II database and consists of images of individuals at the three years intervals.

The organization of this chapter is as follows: A mathematical framework for the use of template ageing is presented in Section 5.2. All necessary steps to calculate the transformation matrix relationship are also explored in this section. The template ageing based on linear transformation is also derived in the same section. Two face recognition systems using the feature template ageing are proposed in Section 5.3. Section 5.4 presents the effect of gender and age groups based on the proposed template ageing scheme. Finally, Section 5.5 looks into key summarised remarks of this chapter.

5.2 Proposed scheme

In this section, a modified face recognition system has been introduced. The proposed modification introduces a transformation so that the natural variation between the enrolled and the verification data (due to ageing) is minimised in order to improve the system performance.

5.2.1 System block diagram

A typical face recognition system comprises of four modules such as image acquisition, pre-processing and normalization, feature extraction, and matching. For instance, individuals who are enrolled (based on enrollment stage) present samples of their biometric data to the system which initially involves the acquisition, pre-processing and feature extraction modules. Then extracted features (biometric templates) are stored in the database in order to be ready for authentication purposes. After some period of time when an individual needs to be verified by the biometric system (authentication stage), firstly the individual presents her/his biometric data again to the system which involves the acquisition, processing and feature extraction modules. Finally, extracted features are compared against the stored templates, and the individual's claimed identity would be verified/denied which involves the database and matching modules. The matching module compares the acquired test image with relevant ones in the template database and generate a decision based on the similarity of both images. In the proposed system, an additional module is introduced to artificially age the enrolled templates prior to comparison by the matcher as shown in Figure 5.1. The motivation is to reduce the variation between the enrolled and test data in order to achieve a more reliable matching outcome.

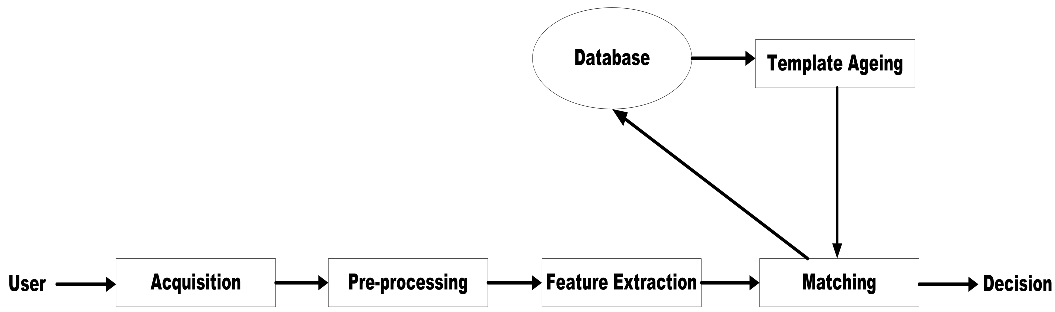

Figure 5.1 Proposed face recognition system with template ageing

5.2.2 Proposed template ageing transformation

Let E be an enrolled template represented by an n-dimensional feature vector $E = e_1, e_2, e_3, \ldots, e_n$ and V be a test data. Let, $V = v_1, v_2, v_3, \ldots, v_n$ which is acquired after a certain time interval in a conventional biometric system..

Both feature vectors, E and V are compared and if the similarity is greater than a pre-set threshold, θ, the test data is accepted as a genuine sample. In other words, let similarity score $\delta = distance(E, V)$

$$\text{if } \delta \geq \theta, \text{ then sample is genuine}$$
$$\text{else } \delta < \theta, \text{ then sample is rejected.}$$

As ageing causes large variation between E and V, there is a high probability that even for a genuine user, δ will fall below the threshold which causes the sample to be rejected.

Suppose there exists a linear mapping $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$, such that, it transforms E in such a way that $\delta_{V,E} < \delta_{V,T(E)}$. The outcome of this linear mapping will reduce the probability that a genuine user will be rejected. Let the artificially aged template $T(E)$ be represented by R.

Conversely, this property allows show to determine how a matrix acts as a mapping. If we have a matrix $A = (a_1, \ldots, a_n)$ where a_i are the column vectors of A, then the action of the mapping induced by multiplication by A will be $Ae_i = a_i$ Multiplying A by the ith standard basis vector has the effect of selecting the ith column of A.

It can clearly see the action of the mapping in this way; the matrix maps the ith standard basis vector to it's ith column.

Let us assume, the variation (due to ageing) between E and V can be represented by a linear transformation such that,

$$V_j = a_{0j} + \sum_{i=1}^{n} a_{ij} e_i \tag{5.1}$$

where a_{ij} are scaler quantities. We can then show v using the relationship as in Equation 5.2.

$$\begin{bmatrix} 1 & e_1 & e_2 & e_3 & \cdots & e_n \end{bmatrix} \begin{bmatrix} a_{00} & a_{01} & a_{02} & \cdots & a_{0j} \\ a_{10} & a_{11} & a_{12} & \cdots & a_{1j} \\ a_{20} & a_{21} & a_{22} & \cdots & a_{2j} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n0} & a_{n2} & a_{n3} & \cdots & a_{nj} \end{bmatrix} = \begin{bmatrix} 1 & v_1 & v_2 & v_3 & \cdots & v_n \end{bmatrix} \tag{5.2}$$

We propose to use this transformation A to artificially age an enrolled template before any comparison.

Now, for identity authentication, let the similarity score $\delta^e = compare(T(E), V)$ and

$$\begin{gathered} \textit{If } \delta^e \geq \theta, \text{ then sample is genuine} \\ \textit{Or else } \delta^e < \theta, \text{ then sample is rejected.} \end{gathered}$$

5.2.3 Transformation parameter estimation

The transformation matrix A can be derived from a set of known samples, using the least square approximation technique, by minimizing the error between the observed and estimated feature values.

$$error\varepsilon^2 = \Sigma(V - a_{0i} - \Sigma a_{ij} E_i)^2 \tag{5.3}$$

where a_{aj} should be such chosen that ε^2 is minimum.

Hence, by setting $\frac{\partial \varepsilon^2}{\partial a_{ij}} = 0$ we can then deduce the transformation matrix A by solving the systems of linear equations shown in Equation 5.4.

$$\begin{bmatrix} N & \Sigma e_1 & \Sigma e_2 & \Sigma e_3 & \cdots & \Sigma e_n \\ \Sigma e_1 & \Sigma e_1^2 & \Sigma e_1 e_2 & \Sigma e_1 e_3 & \cdots & \Sigma e_1 e_n \\ \Sigma e_2 & \Sigma e_1 e_2 & \Sigma e_2^2 & \Sigma e_2 e_3 & \cdots & \Sigma e_2 e_n \\ \Sigma e_3 & \Sigma e_1 e_3 & \Sigma e_2 e_3 & \Sigma e_3^2 & \cdots & \Sigma e_3 e_n \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \Sigma e_n & \Sigma e_1 e_n & \Sigma e_2 e_n & \Sigma e_3 e_n & \cdots & \Sigma e_n^2 \end{bmatrix} \begin{bmatrix} a_{0j} \\ a_{1j} \\ a_{2j} \\ a_{3j} \\ \vdots \\ a_{nj} \end{bmatrix} = \begin{bmatrix} \Sigma v_j \\ \Sigma e_1 v_j \\ \Sigma e_2 v_j \\ \Sigma e_3 v_j \\ \vdots \\ \Sigma e_n v_j \end{bmatrix} \tag{5.4}$$

5.2.4 Feature vector partition

In order to obtain A, there should be m known sample pairs of E and V, where $m \geq n+1$. There may be cases, where the feature vector dimension is quite large and there may not be sufficient known samples available to obtain T. In such cases, we may split the feature vector into k partitions such that each partitions will have its own A and combined all the $A's$ into a feature vector . This is shown in Figure 5.2.

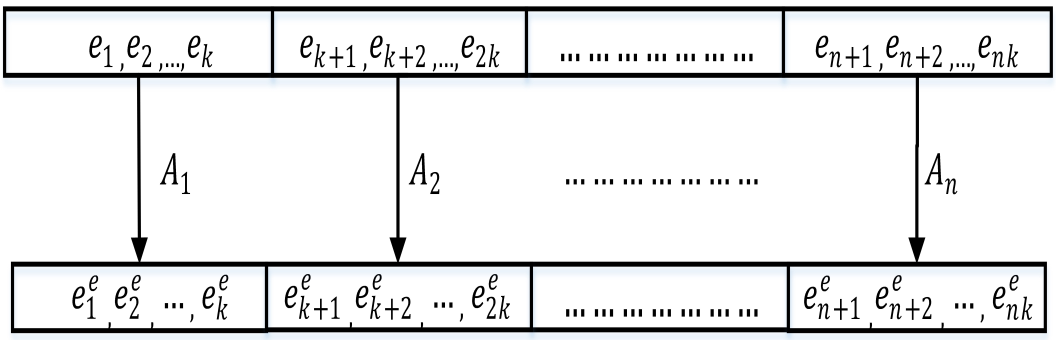

Figure 5.2 Feature vector partition

5.3 Experimental setup

There are two different face recognition algorithms implemented to explore the performance of the proposed template ageing technique. These are Gabor Principal Component Analysis (GaborPCA) and Local Pattern Patterns (LBP).

For the first system, Gabor filters [129] were used for feature extraction in the face recognition scheme. A set of 40 Gabor filters with 8 different frequencies and 5 orientations were used for the feature extraction. The feature vectors consist of the Gabor magnitudes extracted and resulting feature dimension is where training samples are limited, it is difficult to model class distributions well if the feature vectors are too large. Use of a feature reduction scheme can reduce the complexity of the task and hence improve the performance of a system. For this study, Principal Component Analysis (PCA) have been used. PCA uses an orthogonal linear transformation to convert a feature vector (possibly with correlated elements) into a set of uncorrelated values in a lower dimensional space. More details of PCA technique can be found in Section 3.5.1.

For the GaborPCA algorithm, the Euclidean distance between the test data V and the enrollment data $E(orT(E))$ are used as given by equation 5.5.

$$euc_n = \sqrt{(V_1 - T(E)_1)^2 + (V_2 - T(E)_2)^2, \cdots, (V_n - T(E)_2^n)^2} \tag{5.5}$$

where $V = V_1, V_2, \ldots, V_n$ and $T(E) = e_1, e_2, \ldots, e_n$

On the second system, Local Binary Patterns(LBP) are used as the feature. In this particular implementation, the LBP parameters and other settings that have been chosen are as follows:

- LBP operator. In this experiment, the $LBP^{u,2}_{(8,1)}$ operator was used to limit the length of the feature vector.

- Uniform LBP patterns. The total feature vector for an image contains $k^2(P(P-1)+3)$ bins.

- Face regions. The images are divided into k regions; as such the length of the feature vector becomes k times larger, in this experiments the images are divided into 49 regions.

- Region weights. As shown in the Figure 3.15 in Section 3.4.4, regions with mainly background pixels are excluded out by assigning them zero weights. The eye and mouth-regions are assigned with the highest weight, because in the numerous trial and testing, it was found that the eyes and mouth are probably the most important characteristic in the ageing for a person. This is the amended weighted LBP used for this study after running numerous experiments to gain the best results for the effect of ageing in face recognition system.

For an image which was divided into 49 regions and eight sampling points on the circles, the feature vector has a size of 2891 bins.

Utilising the LBP algorithm, the χ^2 distance is used to compare between the test data V and aged template data $T(E)$ as shown in Equation 5.6.

$$\chi^2_n = \sum \frac{(V_1 - T(E)_1)^2}{T(E)_1}, \ldots, \sum \frac{(V_n - T(E)_n)^2}{T(E)_n} \tag{5.6}$$

where $V = V_1, V_2, \ldots, V_n$ and $T(E) = e_1, e_2, \ldots, e_n$

A new subset of images from the MORPH-II database [15] has been adopted to test the performance of the proposed template ageing method for the two face recognition systems described above. This was deemed necessary as the previous subset does not contain a uniform distribution of gender and age groups; the effect of which is also investigated in the following experiments in this chapter.

As all the experiments here investigates the effect of time lapse on performance, the experiments consisted 160 images where the time interval between enrollment and verification is 3 years are included in this subset. There are also some images which are at 1 year intervals. These were only used to assess the base performance of the systems and these were not used in template ageing process or experiments. 80 subjects were picked with ages between 20-55 years old at their first enrollment. For example, subject 1 with enrollment age of 20 and three years later as the age of 23 were used for testing.

There were 50 males and 30 females, while there were 45 peoples (aged between 33 55 years old (the "mature" group) and 35 subjects were aged between 2032 years old (the "young" group). Table 5.1 shows the demographic breakdown of the dataset.

Table 5.1 Demographic breakdown of the dataset

Description	Mature	Young	
Male	29	21	50
Female	16	14	30
	45	35	80

5.4 Experimental evaluation

Only the verification scenario has been investigated in this study. For comparative analysis, only the Genuine Accept Rate (GAR) and False Accept Rate (FAR) will be reported. In a biometric verification scenario, GAR denotes the portion of the genuines accepted as genuine whereas False Accept rate (FAR) is the fraction of impostors users accepted as genuines. The FAR and GAR depend on the operational threshold. For comparative analysis, GARs at thresholds set where FARs are 0.01 and 0.05 respectively will be utilised.

5.4.1 Effect of the proposed template ageing scheme

The first experiment assesses the effect of ageing as the interval between enrollment and verification images increases. Table 5.2 represents performance degradation of the face recognition systems as the age intervals increases when no template ageing has been corporate. The performances at 1 year interval has also been included for comparison. The verification rates (GAR) presented here are for FAR =0.01 and

FAR= 0.05 values. The notation used in the age intervals explained the present time of enrollment taken for an individual and the same individual after a period of year(s) later. For instance, $(0, 1]$ means the time intervals of subject A taken between present time(enrollment) and 1 year later.

Table 5.2 Verification performance without template ageing (for all the subjects)

	Algorithm / Age Interval	GAR *Gabor-PCA*	GAR *LBP*
FAR =0.01	(0,1]	0.83	0.86
	(0,3]	0.43	0.47
FAR =0.05	(0,1]	0.90	0.90
	(0,3]	0.44	0.67

From the observations, it is evident that as the age interval increases the recognition performance decreases quite significantly. With the GaborPCA algorithm , GAR at 1% FAR after 1 year interval is 0.83 compared to LBP with 0.86. The GAR for both the algorithms decreases to an average of 0.43 and 0.47 respectively (i.e almost half) when the interval between training and testing data increases to 3 years. Similar trend was also noticed when FAR=0.05, of the two schemes LBP algorithm performed relatively better than GaborPCA algorithm. With GaborPCA algorithm, the GAR reduced to half when the time intervals increased from 1 year to 3 years between enrollment and verification.

Table 5.3 Verification performance with template ageing (for all the subjects)

		GAR	
	Description / Algorithm	Without template ageing	With template ageing
FAR =0.01	Gabor-PCA	0.43	0.68
	LBP	0.47	0.62
FAR =0.05	Gabor-PCA	0.44	0.76
	LBP	0.67	0.75

In Table 5.3, the template transformation method are applied in this experiments. In this experiment, all the available image pairs with 3 years interval were first used to obtain the ageing function. Subsequently, the transformation was applied during the test. It can be seen that both the algorithms observed an increment in their performance. For the GaborPCA algorithm, the genuine accept rate increases from 0.43 to 0.68 at FAR=0.01, while for LBP the increment is smaller and increases GAR from 0.47 to 0.62. However, in this experiment, GaborPCA showed a larger increment compared to LBP algorithm. The GaborPCA algorithm has a constant improvement of 0.25 at FAR=0.01 and 0.32 at FAR=0.05.

When observing the FAR =0.05 both algorithms represent a slight improvement when comparing between enrollment and 1 and 3 years age intervals. In the 1 year age interval both algorithms have a GAR of 0.90, when template ageing was in-cooperated the GAR has been noticeably improved with at 0.08 for LBP and 0.32 for GaborPCA algorithms. Eventhough this is for the whole dataset in the testing experiments, it can be observed that GaborPCA algorithm performed better with GAR of 0.07 when the FAR threshold changes from FAR=0.01 to FAR=0.05.

The following experiment further explores the use of a smaller and disjoint set of data for estimation of A. The populations are segmented into disjoint training and test sets and the results are presented in Table 5.4. Each experiment was run 5 times each with random partitions data and the mean GAR and associated standard deviations are reported. The choice of this number of run was chosen, as it was found that after running different set of experiment, 5 times is the recommended method for cross validation purpose.

Table 5.4 Mean GAR for disjoint training and test population for LBP.

			GAR	
FAR	**Algorithm**	No. of training samples	Without template ageing	With template ageing
0.01	*Gabor-PCA*	*40*	0.50 (±0.01)	0.52 (±0.03)
0.05			0.51 (±0.02)	0.54 (±0.05)
0.01		*60*	0.51 (±0.02)	0.53 (±0.04)
0.05			0.59 (±0.03)	0.64 (±0.06)

It can readily be observed that as the training size increases, the system accuracy increases dramatically. This can be noted that in the LBP algorithm, when FAR=0.01, it initially verify from 0.50 to 0.52 in the 40 subjects training set while as the subjects increases to 60 subjects the GAR increases from 0.51 to 0.54. As the FAR increases to 0.05, the GAR for this algorithm were improved from 0.59 to 0.64. This can be compared to 0.54 during FAR =0.01. The gain through the proposed template ageing is rather limited here which may be attributed to the smaller training data but needs further investigation. Although the gain is limited, it can be noted that the standard deviation for each result shows minimum changes (range between 0.01 to 0.06) and thus it can be taken as an average for each results.

Table 5.5 Mean GAR for disjoint training and test population for Gabor-PCA.

			GAR	
FAR	**Algorithm**	**No. of training samples**	Without template ageing	With template ageing
0.01	*LBP*	*40*	0.36($\pm$0.02)	0.41 ($\pm$0.03)
0.05			0.38 ($\pm$0.02)	0.44 ($\pm$0.04)
0.01		*60*	0.46 ($\pm$0.03)	0.48 ($\pm$0.04)
0.05			0.54 ($\pm$0.04)	0.57 ($\pm$0.05)

In the case of GaborPCA algorithm as stated in Table 5.5, it can readily be observed that as the subjects in the training size increase, the system accuracy has less favourable results compared to the LBP algorithm. In the FAR=0.01 threshold, it has initially verified that GAR increases from 0.36 to 0.41 in the 40 subjects training set while with 60 subjects the GAR increases from 0.38 to 0.44. As the FAR increases to 0.05, the GAR for both 40 and 60 subjects also improved with an average of 0.09 compared to 0.03 in FAR=0.01. Although, the gain is lower that LBP algorithm, overall GaborPCA algorithm has a more consistent GAR with an average of increment between 0.02 - 0.05 in the FAR = 0.01 compared to larger gaps in LBP.

In conclusion, from Table 5.4 and Table 5.5 above, it can be observed that both algorithms showed an increase in the Genuine Accept Rate. In the FAR=0.01 for 40 subjects, LBP showed an increase of 0.02, however GaborPCA has a better improvement of 0.05 for 50% of the total dataset. Both algorithms show a significant increase at FAR=0.05 when there are 60 subjects (75% of total dataset) in the training dataset with 0.64 and 0.57 respectively.

5.4.2 Effect of gender and age groups on the proposed template ageing scheme

This section carries out a further detailed experiment focusing on gender and age groups. The aim of these experiments was to further investigate whether the facial ageing manifests differently in the population subgroups. For the experiments reported in Table 5.6, the population are split into gender and age group. There are 50 males and 30 females in the gender subgroups while 35 “young”(20 - 32 years old) and 45 “mature”subjects (33 -55 years old) group in the age category. Initially, all the available data was used for the estimation of the transformation matrix A.

From the observation as illustrated Table 5.6, all four categories show a significant increase in system accuracy. The gender category performs slightly better than the age group. In the gender category, male performs slightly better than the female, the GAR improved significantly with an average of 0.39 when template ageing was cooperated into the experiment. Within the gender category, the male shows a double increase GAR of 0.43 . For the female it was initially recorded a GAR of 0.52 without template ageing, after applying the proposed template ageing, the GAR has showed an increase of 0.35.

Table 5.6 Performance for Gender and Age group categories using Gabor-PCA.

	Description		Training Samples	GAR: Without Template Ageing	GAR: With Template Ageing
FAR = 0.05	Gender	Male	*50*	0.45	0.88
		Female	*30*	0.52	0.87
	Age Group	Young	*35*	0.52	0.82
		Mature	*45*	0.47	0.77

The GAR for age group after template ageing recorded an increment of 0.30 for both "young" and "mature" age groups. When looking at the age groups category, both "young" and "mature" , has a constant GAR increment of 0.30. In the "young" age group, the GAR increases to 0.82 while "mature" age group increases to 0.77. Within the age group category, "mature"perform slightly better than the "young" age group of a difference GAR of 0.05.

This can be noted that in the GaborPCA algorithm, the best performer for gender category is the male gender group with a GAR of 0.88 whereas from the age group category "mature" age group has an initial verification rate of 0.47 to 0.77 for the all subjects allocated in the training phrase. .

Table 5.7 Performance for Gender and Age group categories using LBP

	Description		Training Samples	GAR: Without Template Ageing	GAR: With Template Ageing
FAR = 0.05	Gender	Male	*50*	0.42	0.98
		Female	*30*	0.36	0.91
	Age Group	Young	*35*	0.35	0.95
		Mature	*45*	0.47	0.97

Table 5.7 shows the results obtained from the experiments based on the LBP algorithm. This can also be noted that in the LBP performed consistently well for all categories with an average GAR of 0.95. The best performer for the LBP algorithm in the gender category is the male gender with the GAR difference of 0.56. Eventhough the female shows a lower GAR when compared to male gender, but in terms of improvement after applying the proposed template ageing, both genders captured similar increment of GAR. It is also noted that both "male" and "female" increased more than double from initial GAR with template ageing.

Within age group category, "mature" age group is the best performer with an improvement GAR of 0.50. In theory, the "mature" age group should perform better than "young" age group as the facial features in the older cohort are more notable and prominent than the younger peer group. In this experiment, although the "young" age group has a lower GAR when compared to "mature", however in terms of difference, "young" shows a triple increment GAR of 0.60.

In conclusion, it can be observed that LBP algorithm performed better than Gabor-PCA algorithm when the overall populations is split into smaller population subgroups. The results show in LBP algorithm has a higher accuracy system of average GAR of 0.95. It is also derived that LBP algorithm performs better in a small segmentation population as compared to the whole population with the GAR of 0.75 when FAR=0.05 shown in Table 5.3.

In the following experiment, disjoint training and test sets were used for both the gender and age grouping categories. In the male gender group , 29 samples were used in training (remaining 20 for test) while in females there were 16 training samples (and 15 test samples) respectively. Similarly, in the age groups, the "young" category there are 21 training samples (and 15 test samples) while in the mature age group there were 24 samples for training data (remaining 20 for test). All the experiments were carried out for 5 simulation runs with random partition of the data and mean standard deviation of the GAR values are reported here for FAR=0.05.

Table 5.8 Verification performance using disjoint training and test dataset for Gabor-PCA algorithm

				GAR	
	Description		**Training Samples**	**Without Template Ageing**	**With Template Ageing**
FAR = 0.05	Gender	***Male***	*29*	0.59(±0.03)	0.65(±0.02)
		Female	*16*	0.28(±0.02)	0.28(±0.03)
	Age Group	***Young***	*21*	0.31(±0.05)	0.32(±0.06)
		Mature	*24*	0.56(±0.03)	0.63(±0.04)

Table 5.8 shows the corresponding GAR values with and without template ageing for GaborPCA algorithm. Despite the smaller training set, the GAR rate increased noticeably for some of the categories. This can be noted that with the GaborPCA the gain is relatively small when compared to Table 5.9. In the gender category, the male has an increment of 0.06 whereas female have no changes in GAR after implementing the proposed template ageing. On the other hand, the gain by the proposed template ageing is much more obvious. With the exception of the female gender group, all the remaining groups experienced large increase in their GAR values. For the male population the GAR value increases to 0.65. It is also evident that the female gender category performed rather poorly for both the algorithms with or without template ageing. The template ageing, in fact, even decreased in sample sizes, the GAR are still achievable. This may again be attributed to the very small training data set that was available, but most likely this is due to the fact that the ageing process is reflected in a more complex way on the extracted feature than the linear transformation proposed here. But it is evident that the ageing process is different for each of these subcategories as evidenced by their larger gain in GAR when group specific ageing transformation had been introduced. Further gain may be achieved by increasing the granularity of these categorizations even with the proposed method and also by providing more training data.

Table 5.9 tabulated the corresponding GAR values with and without template ageing for LBP algorithm. Due to a smaller dataset all of the categories using LBP algorithm have a lower GAR when testing without the template ageing approach. When applying the template ageing, there is an improvement for all of the categories. In the "mature" age group, it can be seen that there is an increase GAR of 0.13 with 20 testing samples.

Table 5.9 Verification performance using disjoint training and test dataset for LBP algorithm

				GAR	
	Description		**Training Samples**	**Without Template Ageing**	**With Template Ageing**
FAR = 0.05	**Gender**	*Male*	*29*	0.45(±0.05)	0.65(±0.04)
		Female	*16*	0.27(±0.05)	0.32(±0.06)
	Age Group	*Young*	*21*	0.34(±0.04)	0.47(±0.04)
		Mature	*24*	0.43(±0.03)	0.64(±0.03)

Within the age group category, “mature” age group is the best performer with an increase GAR of 0.21 when applying the proposed template ageing approach into the face recognition system. It can be seen that during the division of subjects in Table 5.5, when there was 50% of subjects in the training samples , the GAR showed an improvement from 0.51 to 0.54. Similarly, when applying the comparable scenario of $\approx$ 50% testing samples in the disjoint tarining/testing dataset, the male gender (29 out of 50 subjects) and“ mature” age group (24 out of 45 subjects) performed slightly better in smaller population. The male achieved a GAR of 0.65 while “mature” achieved a GAR of 0.64.

It can be observed that as disjoint training and test sets were used, both GaborPCA and LBP algorithms does not perform well. This is due to the lack of images per subject that are in the training samples. However, it is noted that “male” group for both algorithms observed an increased of GAR when the population are segmented into gender (without template ageing) compared to all population in the training samples being used in the implementation. It can be observed that LBP algorithm performed better than GaborPCA algorithm when the overall population is split into smaller population subgroups.

In conclusion, when comparing between 2 different face recognition algorithms, it can be seen that LBP algorithm has a better accuracy in terms of overall experiments. GaborPCA algorithm performs better in larger population category whereas LBP works better in a smaller population.

5.5 Conclusion

The aim of this study was to investigate the relative performance of two different feature extraction and matching algorithms for three (3) age intervals subjects using transformation matrix method. The experiments used a facial image database containing face images acquired from enrollment age to within 3 years period. It was observed that system performance improved significantly when "male" and "mature" age groupings were introduced and classification schemes were deployed to compare the two distinctive feature selection. It was also testified that for this dataset the GAR at 0.01 and 0.05 FAR was highest for the all the experiments except "female" gender. The highest verification rate for GaborPCA is 0.65 which is male population and second performer is GAR of 0.63 for "mature" age grouping. For the LBP algorithm, the highest verification rate similarly falls under male population with a GAR of 0.65 and "mature" age population with 0.64. Further work will explore how an adaptive system can be designed to harness this age dependency for maximum recognition performance.

In the next chapter a novel template ageing using non-linear transformation will be presented.

CHAPTER 6

Template ageing using non-linear transformation

6.1 Introduction

This chapter presents a novel feature-based template ageing scheme using non-linear transformation technique. Here Neural Networks (NN) are used to capture the underlying ageing transformation between the enrollment and verification samples. In Chapter 5, we investigated the performance degradation in facial recognition due to age interval between the enrolled and the test samples and attended template ageing using a linear transformation system to overcome the problem. As there are limitations in the linear transformation, in general, a linear forward problem will not necessarily be shift-invariant and size of image and data spaces are different. This chapter will explore a similar relationship using an non-linear mapping method. As before, the feature vector from enrollment and feature vectors extracted after a time lapse were used to estimate the transformation matrix which relate the enrollment to the latter age.

The organization of this chapter is as follows: The introduction of Neural Network and associated back-propagation algorithms are explained in Section 6.2. The experimental setup and results presented in Section 6.3. It also includes experimental results when soft biometrics are added. Also multi-classifier schemes are tested in Section 6.3.2. Finally some concluding remarks for the chapter are provided in Section 6.4.

6.2 Neural Network

A Neural Network (NN) is a powerful data modeling tool that is able to capture complex input/output relationship. The neural network technology stemmed to develop an artificial systems. The network is composed of a large number of highly interconnected processing elements, called neurons, working in parallel to solve a specific problem.

Initially, random weights and thresholds are assigned to the network. These weights are updated every iteration in order to minimize the cost function or the mean square error between the output vector and the target vector. Figure 6.1 shows the architecture of a typical feed-forward neural network.

In general, the parameters of a neural network, i.e. the weights and biases, are learned using a training data set. However, as the space of possible weights can be very large and of high dimension the analytical determination of these weights might be very difficult or even infeasible.

There are two principal training modes which determine the way the weights are updated:

- Online training: After presentation of each training example, the error is calculated, and the weights are updated accordingly.

- Offline training: The whole training set is propagated through the NN, and the respective errors are accumulated. Finally, the weights are updated using the

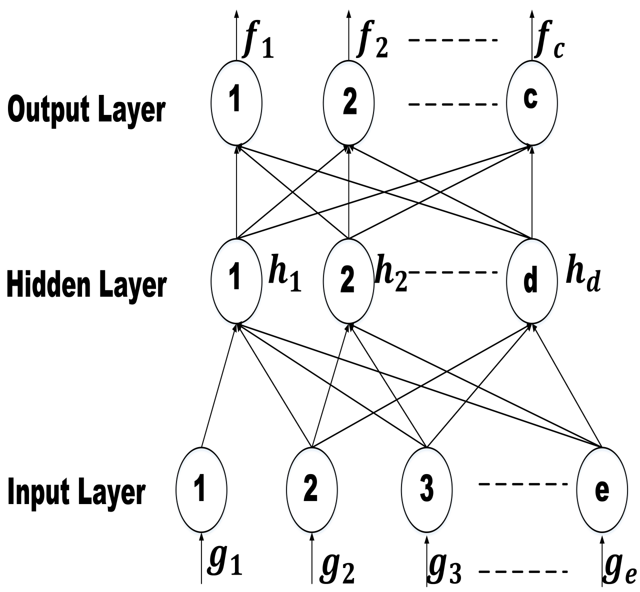

Figure 6.1 Neural network architecture

accumulated error. This is also called batch training.

A successful face recognition methodology depends heavily on the particular choice of the features used by the pattern classifier. The Back-Propagation is the best known and widely used learning algorithm in training Multilayer Perceptrons (MLP) [170]. The MLP refer to the network consisting of a set of sensory units (source nodes) that constitute the input layer, one or more hidden layers of computation nodes, and an output layer of computation nodes. The input signal propagates through the network in a forward direction, from left to right and on a layer-by-layer basis. BPNN are capable of approximating arbitrarily complex decision functions. Back propagation performs a gradient descent within the solution's vector space towards a global minimum.

In the context of neural networks, the BPNN algorithm has initially been presented by Rumelhart et al. [171]. It is a supervised learning algorithm defining an error function and applying the gradient descent technique in the weight space in order to minimize error. The combination of weights leading to a minimum of error is considered to be a solution of the learning problem. Note that the BP algorithm does not guarantee to find a global minimum which is an inherent problem of gradient descent optimization.

The BPNN algorithm though which a neural network learn the input-output relationship, involves two phases. During the first phase, the input vector is presented and propagated forward through the network to compute the output values f_c for each output unit. This output is compared with its desired value, resulting in an error signal δ_i for each output unit. The second phase involves a backward pass through the network during which the error signal is passed to each unit in the network and appropriate weight changes are calculated.

Input for hidden layer is given by

$$net_d = \sum_{z=1}^{e} g_z w_{dz} \tag{6.1}$$

The units of output vector of hidden layer after passing through the activation function are given by

$$h_d = \frac{1}{1 + exp(-net_d)} \tag{6.2}$$

In the same manner, input for output layer is given by

$$net_c = \sum_{d}^{z=1} h_z w_c z \tag{6.3}$$

and the units of the output vector of output layer are given by

$$f_c = \frac{1}{1 + exp(-net_c)} \tag{6.4}$$

Other activation functions e.g. hyperbolic tangent activation may also be employed depending on the range of the desired output values. The net error can be calculated by,

$$E = \frac{1}{2} \sum_{i=1}^{d} (f_i - t_i)^2 \tag{6.5}$$

If the error is less than a predefined limit, training process will stop; otherwise weights need to be updated.

For weights between hidden layer and output layer, the change in weights is given by

$$\Delta w_{ij} = \alpha \delta_i h_j \tag{6.6}$$

where α is learning rate coefficient that is restricted to the range [0.01, 1.0]. The learning coefficient α controls the size of a step against the direction of the gradient. If α is too small, learning is slow; if too large, the process of the error minimization can be oscillatory. Here, h_j is the output of neuron j in the hidden layer and δ can be obtained by

$$\delta_i = (t_i - f_i) f_i (l - f_i) \tag{6.7}$$

where f_i and t_i represents the real output and target output at neuron i in the output layer respectively. Similarly, the change of the weights between hidden layer and input layer is given by

$$\Delta w_{ij} = \beta \delta_{Hi} g_j \tag{6.8}$$

where β is a training rate coefficient that is restricted to the range [0.01,1.0], g_j is the output of neuron j in the input layer. δ_{Hi} can be obtained by

$$\delta_{Hi} = g_i (l - g_i) \sum_{j=1}^{c} \delta_j w_{ij} \tag{6.9}$$

After calculating the weight change in all layers, the weights can be simply updated by

$$w_{ij}(new) = w_{ij}(old) + \Delta w_{ij} \tag{6.10}$$

Updating hidden units process is repeated for each instance in the training set until the error for the entire system is acceptably low, or the predefined number of iterations is reached. To effectively increase the learning rate is to modify the delta rule by including a momentum term.

$$\Delta w(N+1) = d\Delta w(N) - \alpha \nabla E(w(N)) \tag{6.11}$$

where d is a positive constant,$0 \leq d < 0.9$, termed the momentum constant and this is called the generalized delta rule.

Feedforward networks often have one or more hidden layers of sigmoid neurons followed by an output layer of linear neurons. Multiple layers of neurons with nonlinear transfer functions allow the network to learn nonlinear and linear relationships between input and output vectors.

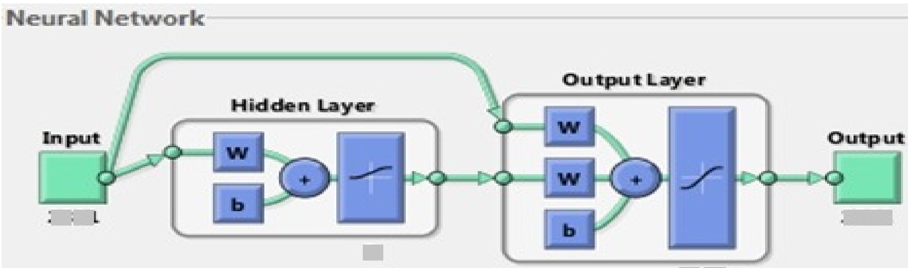

Figure 6.2 Cascade Forward Back propagation Network

CFBPNN models shown in Figure 6.2 and Figure 6.3 are similar to feed-forward networks, but include a weight connection from the input to each layer as well as from each layer to the successive layers. For example, a three layer network has connections from layer 1 to layer 2, layer 2 to layer 3, and layer 1 to layer 3. The three-layer network also has connections from the input to all three layers. The additional connections usually improve the speed at which the network learns the desired relationship [172]. CF artificial intelligence model is similar to feedforward backpropagation neural network in using the backpropagation algorithm for weights updating. Tanh-sigmoid transfer function, log - sigmoid transfer functions were used to reach the optimized status [172].

Al-allaf et al [173] used CFBPNN as part of their face recognition implementation but it did not perform well when compared to other neural network. In this research, we will further explore this model for face ageing purpose. CFBPNN was commonly been used in iris recognition before this [174, 175].

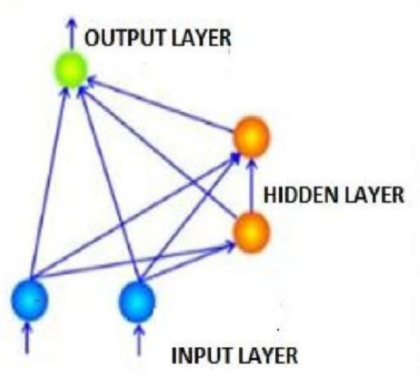

Figure 6.3 Cascade Forward Back propagation Network

6.3 Experimental setup

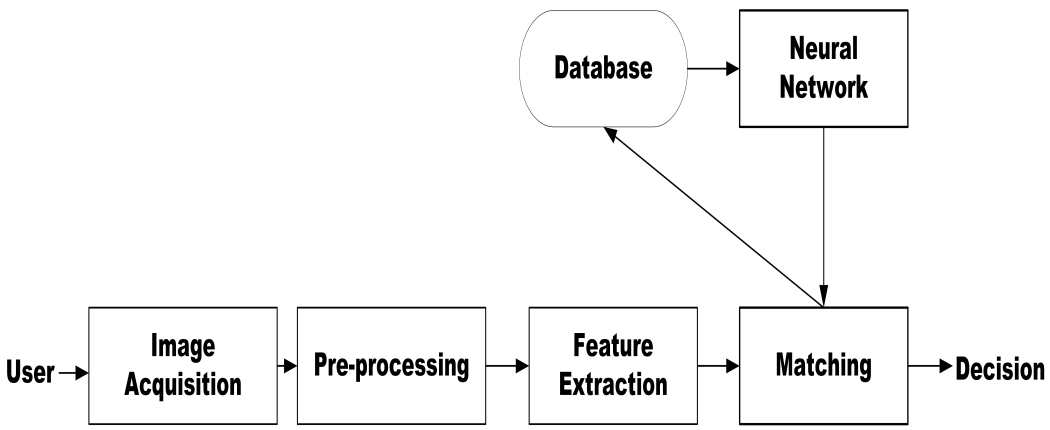

Figure 6.4 Proposed template ageing using back-propagation algorithm

In the proposed system, the Neural Network module acts as a function to artificially age the enrolled templates and predict an estimation of the time lapse template in the output as shown in Figure 6.4. The motivation is to reduce the variation between enrolled and test data using the inherent non-linear transformation of the neural networks.

This study explores both the Feedforward backpropagration (FFBPNN) and Cascade forward backpropagation (CFBPNN) artificial intelligence models. They are trained using the MORPH-II subset which consisted of face images acquired at 3 years intervals. Five different optimization ANN training algorithms were used to train the two models separately to identify the model with the best results for the face recognition system:

- Resilient backpropagation (TRAINRP)
- Conjugate gradient backpropagation with Fletcher-Reeves(TRAINCGF)
- Gradient descent with momentum (TRAINGDM)
- Conjugate gradient backpropagation with Powell-Beale (TRAINCGB)
- Basic gradient descent (TRAINGD)

For comparative analysis, we only reported the Genuine Accept Rate (GAR) and False Accept Rate (FAR). We will be reporting, for our comparative analysis here, GARs at thresholds set where FAR is 0.05. The number of nodes in input layer is equal to the dimension of the feature vector incorporating the GaborPCA or LBP features.

6.3.1 Experimental evaluation

Various parameters assumed for the network are as follows:

- No. of Input unit = LBP/ GaborPCA feature vector
- Training Function = Traingdm/traingd/ traincgf/traincgb/trainrp
- No. of Output unit = LBP/ GaborPCA feature vector

- Hidden layer Transfer Function = log-sigmoid
- Output layer Transfer Function = tanh-sigmoid
- Learning Function = Least mean squared (LMS)
- No. of epoch = 1000
- Optimum value of goal = 0.01
- Ratio for training, validation,testing = 0.70, 0.15, 0.15
- Training Set: this data set is used to adjust the weights on the neural network.
- Validation Set: this data set is used to minimize overfitting.
- Testing Set: this data set is used only for testing the final solution in order to confirm the actual predictive power of the network.

The FFBPNN and CFBPNN trained present fast convergence and the training process terminated within 1000 epochs, with summed squared error (SSE) reaching the pre-specified goal. In the transfer function, log-sigmoid and tanh-sigmoid function are used at all neurons in hidden layer and output layer respectively. The output of Gabor wavelet is an image vector of size 10240 and the output of LBP is an image vector of size 2891.

Comparison of GaborPCA and LBP based schemes for all populations were already presented in Table 5.3. It can be observed that without the template ageing the GAR is quite low. With linear template ageing the performance improved improved noticeably and the GaborPCA performed slightly better than LBP feature vectors

Table 6.1 shows how the performance improves further when a non-linear template ageing is included in the scheme. From the observation, CFBPNN performed slightly better than FFBPNN in both face recognition algorithm. It can be noted that the time taken for CFBPPN to train and test the ANN has an average of 10 seconds faster than FFBPNN algorithm. In related to BPNN algorithms, LBP performed significantly better than GaborPCA. It can be observed that both GAR has a difference of 0.10 when comparing between both algorithms.

Table 6.1 Comparison of GAR for GaborPCA and LBP

Algorithm	GAR at FAR=0.05					
	No Ageing	Linear Ageing	FFBPNN	Time in sec	CFBPNN	Time in sec
GaborPCA	0.44	0.76	0.78	57	0.80	51
LBP	0.67	0.75	0.88	61	0.90	47

Figure 6.5 shows the variation of GAR when varying the number of neurons in the hidden layer with GaborPCA and LBP characterization. In order to show the importance of processing elements,the BPNN was trained with variable hidden units from 2 - 300. For a large number of neurons (20 - 300) in the hidden layer, it was observed that it produced a large error, so accuracy is low. The BPNN generalize poorly. After 10 neurons, error came back to the levels of a system with only 3 neurons in the hidden layer. When there are too many neurons, poor performance is a direct effect of overfitting. The system overfits the training data and does not perform well on novel patterns.

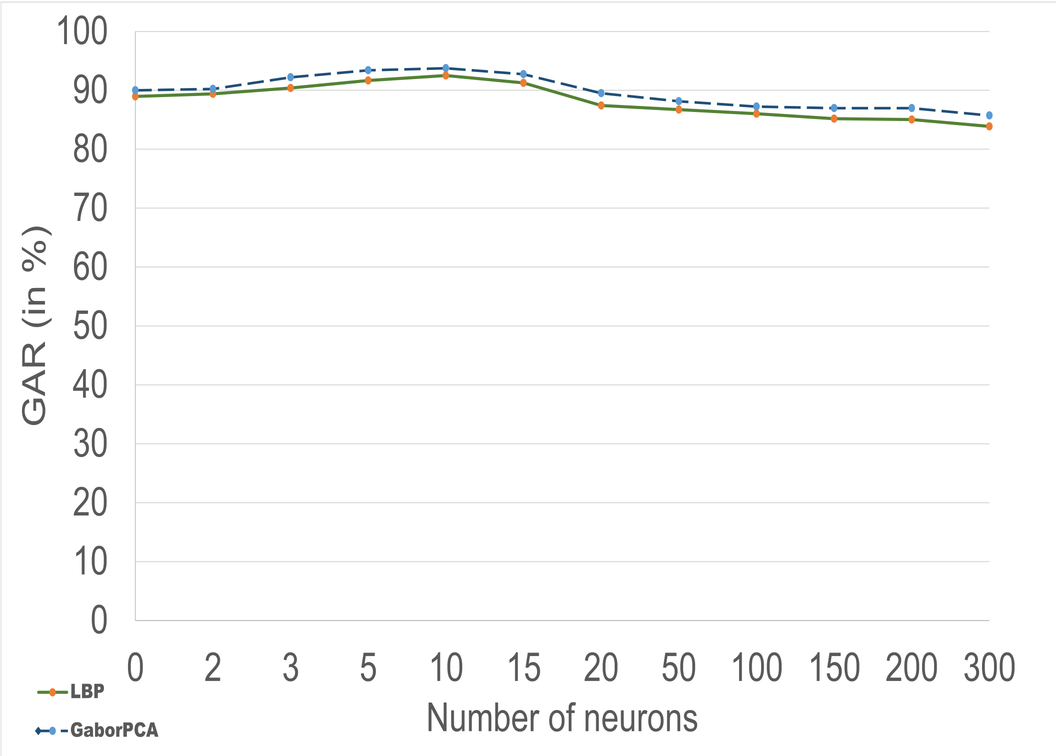

Figure 6.5 GAR with GaborPCA and LBP for BPNN algorithm

Once the initial experiments were carried out using both GaborPCA and LBP features, it was decided to carry out further detailed experiments incorporating on soft biometrics. The aim of these experiments was to further investigate to see if the performance of the methods could be improved when soft biometric information is available.

In Table 6.2, the subjects were split into gender and age group categories. There are 50 males and 30 females in the gender group while 35 subjects in the "young" and 45 subjects in the "mature" age group. Each experiment was run 5 times each with random partitions data and the mean GAR and associated standard deviations are reported.

From Table 6.2, it can be noticed that there are significant improvement when BPNN is implemented for both the feature schemes. The increments are obvious in the "male" and "mature" categories with about 50% increase when compared to no template ageing. In the gender grouping, both "male" and "female" showed an improvement when comparing to linear template ageing in Table 5.8. The "female' groups showed a GAR increase of 0.64 compared to linear template ageing with GAR of 0.28. This can be evidently proved that "male" group are more stable in their facial features and it has a significant consistent improvement.

Table 6.2 Comparison of GAR for Gabor-PCA with FFBPNN.

Biometric Feature	GAR at FAR=0.05				
	FFBPNN for Gabor-PCA				
	TRAINRP	TRAINGDM	TRAINCGF	TRAINCGB	TRAINGD
All	0.85(±0.04)	**0.88**(±0.03)	0.83(±0.03)	0.83(±0.03)	0.86(±0.03)
Male	**0.96**(±0.03)	0.85(±0.02)	0.80(±0.04)	0.80(±0.02)	0.85(±0.03)
Female	0.88(±0.03)	**0.92**(±0.04)	0.83(±0.03)	0.82 (±0.03)	0.81(±0.02)
Young	**0.89**(±0.02)	0.83(±0.03)	0.74(±0.04)	0.75(±0.04)	0.82(±0.03)
Mature	**0.94**(±0.03)	0.93(±0.04)	0.85(±0.01)	0.84(±0.03)	0.87(±0.02)

When looking at the age groups category, the "young" group has increased GAR by 0.57 (in TRAINRP function) compared to linear template ageing. This can be seen that after applying a non-linear template ageing, all of the groups performed favourably in GaborPCA algorithm. Overall, the 2 best performers are from TRAINRP and TRAINGDM.

Table 6.3 Comparison of GAR for LBP with FFBPNN.

Biometric Feature	GAR at FAR=0.05				
	FFBPNN for LBP				
	TRAINRP	TRAINGDM	TRAINCGF	TRAINCGB	TRAINGD
All	0.90(±0.03)	0.94(±0.03)	0.95(±0.04)	0.94(±0.03)	**0.95**(±0.02)
Male	0.92(±0.04)	**0.93**(±0.02)	0.90(±0.02)	0.91(±0.04)	0.93(±0.04)
Female	0.89(±0.03)	0.92(±0.02)	0.92(±0.04)	0.90(±0.02)	**0.93**(±0.03)
Young	0.83(±0.02)	0.84(±0.04)	0.84(±0.03)	0.88(±0.03)	**0.90**(±0.03)
Mature	0.90(±0.04)	**0.94**(±0.03)	0.93(±0.03)	0.92(±0.02)	0.94(±0.04)

For FFBPNN algorithm with LBP feature, Table 6.3 presents the 5 different types of training functions when compared to soft biometric scenario. It can be noted that "female" group preform better in LBP compared to GaborPCA using FFBPNN algorithm. The two best performers associated with TRAINGDM and TRAINGD training function. When comparing with GaborPCA, both"young" and "mature" in the age groups performed slightly better in LBP. It is interesting to note that grouping using soft biometrics achieved a high GAR accuracy of 0.93 on average.

Table 6.4 Comparison of GAR for LBP CFBPNN.

Biometric Feature	GAR at FAR=0.05				
	CFBPNN for LBP				
	TRAINRP	TRAINGDM	TRAINCGF	TRAINCGB	TRAINGD
All	**0.90**(±0.04)	0.83(±0.01)	0.89(±0.01)	0.88(±0.01)	0.80 (±0.01)
Male	0.93(±0.03)	0.84(±0.31)	0.82(±0.04)	0.85(±0.02)	**0.94**(±0.03)
Female	0.92(±0.01)	0.93(±0.04)	0.93(±0.03)	0.89(±0.03)	**0.93**(±0.01)
Young	0.71(±0.02)	0.56(±0.02)	**0.85**(±0.01)	0.71(±0.04)	0.84(±0.01)
Mature	**0.92**(±0.01)	0.79(±0.04)	0.83(±0.01)	0.88 (±0.01)	0.83(±0.02)

Since LBP algorithm performed better in BPNN models, the following experiment shown in Table 6.4 further explores LBP in CFBPNN models. From the observation, only the "mature" age group perform better when compared to FFBPNN. However, it can be observed that all of the age grouping perform better when compared to GaborPCA FFBPNN. It is also noted that it takes a resonably shorter time to run the experiment in CFBPNN compared to FFBPNN.

6.3.2 Multi-Classifier Fusion

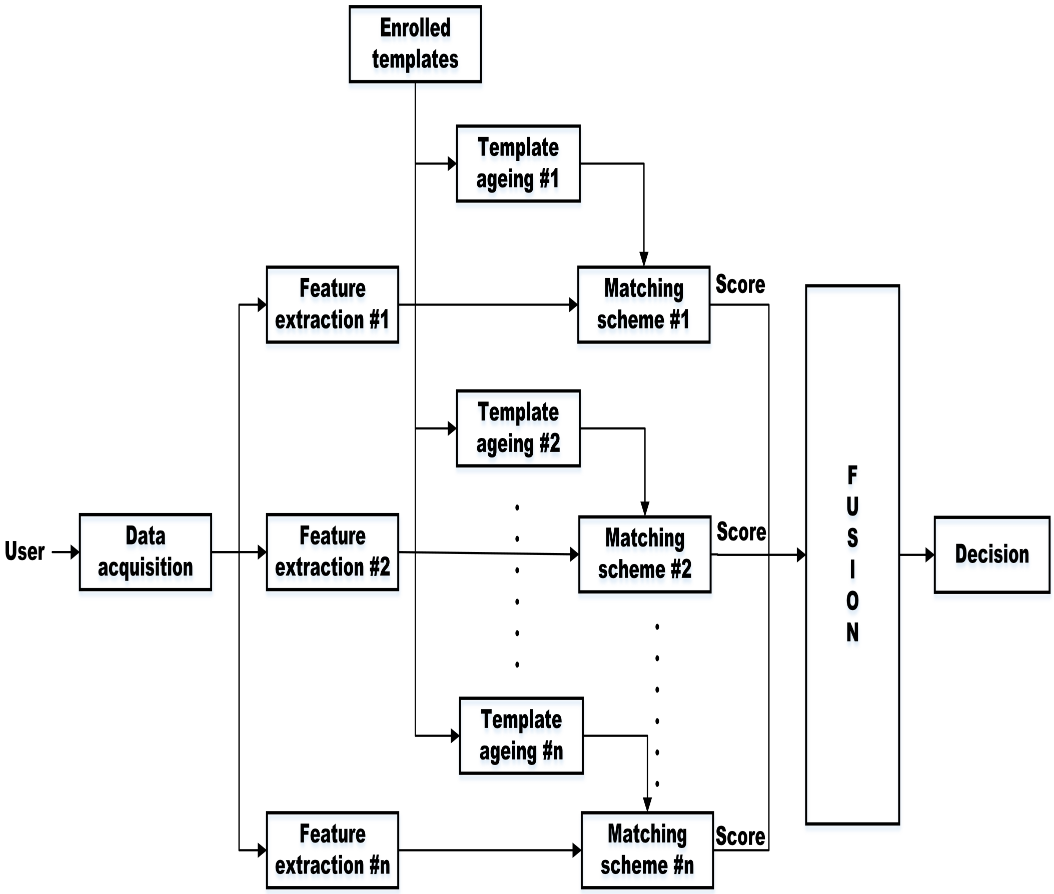

Figure 6.6 A multi-classifiers scheme using different template ageing scheme

It was evident so far that different BPNN schemes performed better for different soft biometric with different settings. The following experiments shall explore whether a multi-classifier approach can produce even better accuracy rates from the template ageing scheme proposed. In the multi-classifier fusion scheme, features extracted from each image are compared to stored template aged by different BPNN schemes and the resulting scores are combined for the final decision. The scheme is illustrated in Figure 6.6. In this section, there will be various experiments to show the effect of face ageing when different types of BPNN algorithms are combined in each experiments.

We only explored the fusion of the superior performing training functions and soft biometric sub-groupings in these experiments.

Table 6.5 Fusion for LBP and GaborPCA FFBPNN

			GAR at FAR=0.05		
Features	Biometric Feature	Fusion Rule	TRAINRP	TRAINGDM	TRAINGD
GaborPCA + LBP	Male	Sum	**0.97**($\pm$0.01)	0.94($\pm$0.01)	0.93($\pm$0.03)
		Product	0.89($\pm$0.03)	0.92($\pm$0.03)	0.89($\pm$0.03)
	Female	Sum	0.93($\pm$0.05)	**0.93**($\pm$0.01)	0.91($\pm$0.02)
		Product	0.89($\pm$0.04)	0.86($\pm$0.02)	0.88($\pm$0.03)
	Young	Sum	0.82($\pm$0.02)	**0.89**($\pm$0.01)	0.84($\pm$0.01)
		Product	0.79($\pm$0.01)	0.74($\pm$0.03)	0.72($\pm$0.03)
	Mature	Sum	0.94($\pm$0.03)	**0.95**($\pm$0.03)	0.93($\pm$0.04)
		Product	0.91($\pm$0.03)	0.92($\pm$0.04)	0.91($\pm$0.02)

Table 6.5 shows the GAR (at FAR=0.05) obtained from combining scores from two classifiers for different BPNN training algorithm. The first classifier uses GaborPCA and second uses LBP. For fusion, both sum and product rule were investigated. Here only the FFBPNN were used. The highest GARs obtained are highlighted in bold.

From the observation, TRAINRP and TRAINGDM were the best performers for sum rule. It can be observed that all soft biometric sub-grouping produced a noticeable higher GAR values when compared to classifiers with no fusion as shown in Table 6.2 and Table 6.3.

Table 6.6 Fusion results from common feature vector and BPNN with 2 training functions

			GAR at FAR=0.05		
Training Function	Biometric Feature	Fusion Rule	GaborPCA FFBPNN	LBP FFBPNN	LBP CFBPNN
TRAINRP TRAINCGF	Male	Sum	0.79(±0.02)	0.92(±0.03)	0.90(±0.03)
		Product	0.80(±0.03)	**0.93**(±0.02)	0.89(±0.03)
	Female	Sum	0.78(±0.03)	**0.89**(±0.03)	0.78(±0.02)
		Product	0.79(±0.04)	0.87(±0.01)	0.81(±0.03)
	Young	Sum	0.79(±0.04)	**0.85**(±0.01)	0.83(±0.03)
		Product	0.78(±0.03)	0.84(±0.03)	0.84(±0.03)
	Mature	Sum	0.81(±0.03)	**0.90**(±0.03)	0.81(±0.02)
		Product	0.82(±0.03)	0.89(±0.02)	0.82(±0.02)

In Table 6.6, the multi-classifier scheme used a common feature vector and one single type of neural network (BPNN) architecture but different training function were incorporated. Again both sum and product rule were explored. It can be readily observed that TRAINRP and TRAINCGF training functions did not perform as good as the other 2 training functions in Figure 6.7. However, it can be observed that LBP FFBPNN algorithm is the best performer for all age groupings. This can be noted that both sum and product rule. All 4 sub-grouping (i.e. “male”, “female”, “young” and “old”) performed differently in this training function combinations.

Table 6.7 Fusion results from common feature vector and BPNN with 2 training functions

			GAR at FAR=0.05		
Training Function	Biometric Feature	Fusion Rule	GaborPCA FFBPNN	LBP FFBPNN	LBP CFBPNN
TRAINRP TRAINGDM	Male	Sum	0.79(±0.02)	**0.96**(±0.01)	0.88(±0.04)
		Product	0.79(±0.03)	0.94(±0.02)	0.90(±0.03)
	Female	Sum	0.76(±0.05)	0.92(±0.02)	0.83(±0.03)
		Product	0.79(±0.03)	0.92(±0.03)	**0.93(±0.03)**
	Young	Sum	0.72(±0.03)	0.83(±0.03)	0.79(±0.03)
		Product	0.74(±0.03)	**0.90**(±0.01)	0.82(±0.04)
	Mature	Sum	0.85(±0.04)	0.90(±0.04)	0.87(±0.02)
		Product	0.86(±0.03)	**0.92**(±0.01)	0.89(±0.03)

Table 6.7 presented similar multi-classifier scenario Table 6.6. This experiment further explore the fusion of 2 different training functions i.e. TRAINRP and TRAINGDM using both sum and product rules. From the observation, LBP FFBPNN and LBP CFBPNN are the best performers for both sum and product rules. When comparing to no classifier scheme for BPNN algorithm in Table 6.3 and Table 6.4, it can be observed that only “male” gender group illustrated a noticeable higher GAR values. For the sub-group of “female” and “young” the GAR value remains the same. However, this experiments presents a higher GAR when compared to Table 6.6.

The following experiment presented the multi-classifier scheme using the common feature vector and one single type of neural network (BPNN) but with 3 training functions were incorporated. Table 6.8 shows the GAR score fusion for the 3 training functions when combined with 3 BPNN algorithms for both GaborPCA and LBP feature approaches. In the gender category, it can be seen that product rule produced a higher GAR compared to sum rule. The noticeable difference in GAR value is “male” gender group which increased by 2% as compared to single classifier in Table 6.3. In the age groups category, both “young” and “mature” presented a noticeable increase in GAR values. The highest GAR value associated with “mature” age groups is in the sum rule which incorporated in LBP FFBPNN with 98%.

In general, when comparing the proposed multi-classifiers to single classifier using non-linear transformation, the GAR values for multi-classifier table is noticeably improved and it showed a significant improvement in the template ageing system. This is evidently proved in the multi-classifiers fusion, in Table 6.8 the 3 training functions has an overall GAR values increased by 3% in average when compared to Table 6.3 for LBP FFBPNN algorithm.

Table 6.8 Fusion results from common feature vector and BPNN with 3 training functions

			GAR at FAR=0.05		
Training Function	Biometric Feature	Fusion Rule	GaborPCA FFBPNN	LBP FFBPNN	LBP CFBPNN
TRAINRP TRAINGDM TRAINCGF	Male	Sum	0.77(±0.01)	0.94(±0.03)	0.92(±0.03)
		Product	0.76(±0.03)	**0.97**(±0.01)	0.93(±0.02)
	Female	Sum	0.76(±0.02)	0.92(±0.03)	0.83(±0.03)
		Product	0.76(±0.03)	**0.94**(±0.01)	0.91(±0.02)
	Young	Sum	0.79(±0.02)	0.93(±0.04)	0.84(±0.05)
		Product	0.79(±0.04)	**0.93**(±0.02)	0.85(±0.03)
	Mature	Sum	0.90(±0.04)	0.95(±0.01)	0.90(±0.02)
		Product	0.89(±0.04)	**0.98**(±0.01)	0.91(±0.03)

It can be readily observed that LBP FFBPNN has a favorable results compared to GaborPCA FFBPNN and LBP CFBPNN algorithms. In conclusion, LBP FFBPNN is the best performers when soft biometric subset is incorporated into the algorithm.

6.4 Conclusion

For both the BPNN algorithms, it can be seen that both feature approaches have a higher accuracy when compared to the linear transformation method. In the single classifier experiments, the GAR are about 90 % accuracy rate for all the sub-categories. This non-linear method has a reduced execution time which is the main advantages of Neural Network based template ageing scenario. Neural Networks are more flexible for solving non-linear tasks. As gradient based method is applied, some inherent problems such as slow convergence and escaping from local minima are sometimes encountered here. But this only affects the training/learning phase of the implementation.

It can be observed that, when BPNN algorithms are applied in similar scenario as in Chapter 5, all of the age grouping showed an increased of about 50% in both GaborPCA and LBP feature approaches. It is also noted that, in general, LBP worked better in the non-linear method while GaborPCA was better in linear transformation schemes.

The next chapter provides the summary, conclusion and recommendations for future work.

CHAPTER 7

Conclusion

7.1 Introduction

In this chapter, a summary of the work covered in this thesis is provided. This will be followed by a discussion of the major findings of this work and suggestions for future research directions. Section 7.2 will lists the contributions to wider research field which have been made by the thesis and briefly conclude the findings from the experiments reported in this thesis. Finally, Section 7.3 discusses some possible future research directions. [155] [156] [157] [158]

The next sections highlights the significant contributions made to the wider field of biometrics by the work reported in this thesis, as well as providing some remarks and directions for future work for researcher in this field.

7.2 Main contribution

The work presented here explores the effect of face ageing, both lateral and longitudinal, on facial biometric system. Reported studies related to template ageing effects in other modalities have agreed that there is generally a considerable reduction of sample similarity across the time between enrolled and subsequently verified biometric data. In this thesis, we investigated various types of ageing issues to improve and provide accurate measure to discriminate between genuine and imposter. Information from different feature extraction approaches have been combined with novel template ageing techniques in a fusion framework and evaluated for a number of verification scenarios. The multi-classifier approach, combining information from various algorithms using fusion provided the best results and was seen to compare well with the state of the art showing the potential effectiveness and viability of this approach. This thesis also analyse and investigate the effect of ageing in biometric system more broadly and extensively on various factor such as gender and age grouping and clearly define and provide insights into how to manage the effect of ageing process for improved reliability of biometric systems.

The main contributions of this thesis are as follows:

- A comparative study with extensive experimentations how biometric system performance varies for population subgroups especially age sub-groups (lateral ageing). This detailed an investigation and analysed the effects of choosing different age bands to adopt in trying to improve the performance of the biometric system. The aim has been to identify the impact of age grouping in the long-term intervals of 0-5 years. Additionally, this study presented an in-depth comparison of using Gabor as the feature extractor with various projection technique variants in order to determine the performance variation. This proposed approach have been used as a guide for any subsequent age related analysis for data processing.

- A comparative study with proposition of a multi-classifier framework in lateral ageing. The framework provides the results of combining various scheme using rule based fusion . The results suggest that such combination may be used to improve the performance of the face ageing biometric systems. The score fusion based systems performed better than the single classifiers.

- An amended weighted LBP scheme for feature extraction. The weighted LBP scheme was introduced in Chapter 5 for improving the performance of longitudinal ageing.

- The proposition of two novels template ageing schemes to deal with long time lapses between enrollment and verification (longitudinal ageing). Chapter 5 presented a novel template ageing scheme using linear transformation to improve the performance of biometric system to overcome longitudinal ageing. This was based on the assumption that there maybe a linear relationship between the feature vectors of time lapsed face images. If this is true and such relation can be established. The verification image can be compared with artificially aged templates for more reliable matching. The approach were extensively investigated using two face recognition systems with a subset of MORPH-II database. Although only 3 years interval data is used in this study, the idea can be extended to any age intervals provided appropriate data are available. The template ageing schemes also explores the gender and age grouping scenarios to be compared with the whole population.

- Proposition of a multi-classifier framework to further boost system performance in the longitudinal ageing scenario. The limited achievements from linear mapping encourages the introduction of complex non-linear mapping. Chapter 6 provides details of novel non-linear template ageing using neural networks. Here we explored the performance results for two different models of BPNN using template ageing approach. Both face recognition are being used for experiment. Significant improved performances were achieved. The fusion of several ageing adjusted approaches were incorporated in a multi-classifier configuration. The final results were very promising.

7.3 Recommendation of future works

The research presented in this thesis has aimed to explore and investigate the facial ageing impact in a biometric systems' performance. Some possible new research ideas based on these issues have emerged from the presented contributions in the thesis which are as follows:

A key issue related to development of ageing methodologies, is the issue of performance evaluation protocols. Another issue that support the idea of evaluation is the availability of dedicated face database. Ideally the database should pair with the same age separated images where age separated faces must be captured under identical imaging conditions so that ageing variation is isolated form other types of variations, allowing researcher to focus their focus on the problem of age progression.

Future research may also explore extraction of different novel face features such as hairline and wrinkle especially in the older age groups. This can be further followed by introducing non-linear transformation approaches such as deep learning model in face ageing biometric systems.

Considering the experience obtained from study on lateral ageing, some aspects such as younger population affect the overall performance. This particular age groups can be further explored on how to minimise the error rate in the biometric system for teenage population. Eventhough, the experiment looked into gender categories, a possible scenario that can be look at are the pose variation and ethnicity. A combination of classification and regression can also be considered to develop a hybrid system.

The experiments carried out in this study only looked at template ageing for a fixed 3 year time lapse between enrollment and verification. Ideally such a system should be capable of dealing with any or variable time lapses and a scheme that takes desire time gap as input would have many useful application.

In this thesis, two dimensional and static gray scale database images are used, but in future, three dimensional and time-varying colour images may be recognized by using real-time setting since ageing could be address more efficiently using 3D models. Given recent technology advances in smart phones, it is also important to generate 3D ageing datasets enabling a more systematic experimentation in 3D face ageing.

The proposed template ageing approaches are generic and may be implemented and fine tuned for other biometric modalities.

List of Publication

Articles in conference proceedings

1. Hayati Yassin, Sanaul Hoque and Farzin Deravi, "*Age sensitivity of face recognition algorithms*," In Emerging Security Technologies (EST), 2013, IEEE Fourth International Conference, pp.12-15, 2013. DOI:10.1109/EST.2013.8

2. Hayati Yassin, Sanaul Hoque and Farzin Deravi, "*Face recognition across ages*," In Brunei International Conference on Engineering and Technology (BICET) , 2016, IET Sixth International Conference (Accepted in July 2016)

Bibliography

[1] A. M. Albert, K. Ricanek, and E. Patterson, "A review of the literature on the ageing adult skull and face: Implications for forensic science research and applications," *Forensic Science International*, vol. 172, no. 1, pp. 1–9, 2007.

[2] J. Grey, "Face recognition using gabor filters, pca and neural networks," 2014. [Online]. Available: http://www.imdb.com/name/nm0000426/

[3] A. P. B. M. RMIP, "Coleen nolan: I dont want to end up looking like this..i must give up smoking," 2009. [Online]. Available: https://aurioleprince.wordpress.com/2009/04/02/coleen-nolan-i-dont-want-to-end-up-looking-like-thisi-must-give-up-smoking/

[4] A. Lanitis, "Comparative evaluation of automatic age-progression methodologies," *EURASIP Journal on Advances in Signal Processing*, vol. 2008, p. 101, 2008.

[5] Y. Fu, G. Guo, and T. S. Huang, "Age synthesis and estimation via faces: A survey," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 32, no. 11, pp. 1955–1976, 2010.

[6] JiwonYu, "Kang min hyuk is all smiles on the set of entertainer," 2016. [Online]. Available: http://www.soompi.com/2016/04/03/kang-min-hyuk-is-all-smiles-on-the-set-of-entertainer/

[7] K. R. Jr and T. Tesafaye, "Morph (craniofacial longitudinal morphological face database)," 2006. [Online]. Available: http://www.faceaginggroup.com/morph/

[8] "Fg-net (face and gesture recognition network)," 2014. [Online]. Available: http://www-prima.inrialpes.fr/FGnet/

[9] V. Struc, R. Gajšek, and N. Pavešić, "Principal gabor filters for face recognition," in *Biometrics: Theory, Applications, and Systems, 2009. BTAS'09. IEEE 3rd International Conference on.* IEEE, 2009, pp. 1–6.

[10] B. Julsing, "Face recognition with local binary patterns," *Research No. SAS008-07, University of Twente, Department of Electrical Engineering, Mathematics & Computer Science (EEMCS)*, 2007.

[11] L. Sánchez López, "Local binary patterns applied to face detection and recognition," 2010.

[12] S. Biswas, G. Aggarwal, N. Ramanathan, and R. Chellappa, "A non-generative approach for face recognition across ageing," in *Biometrics: Theory, Applications and Systems, 2008. BTAS 2008. 2nd IEEE International Conference on.* IEEE, 2008, pp. 1–6.

[13] X. Geng, Z.-H. Zhou, and K. Smith-Miles, "Automatic age estimation based on facial ageing patterns," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 29, no. 12, pp. 2234–2240, 2007.

[14] K. Niinuma, U. Park, and A. K. Jain, "Soft biometric traits for continuous user authentication," *Information Forensics and Security, IEEE Transactions on*, vol. 5, no. 4, pp. 771–780, 2010.

[15] K. Ricanek Jr and T. Tesafaye, "Morph: A longitudinal image database of normal adult age-progression," in *Automatic Face and Gesture Recognition, 2006. FGR 2006. 7th International Conference on.* IEEE, 2006, pp. 341–345.

[16] K. Ricanek, "Facial ageing update and facial modeling," 2012, iSIS Institute-Face Ageing Group, University of North Carolina Wilmington. Last Accessed: November 2014.

[17] A. Lanitis, "A survey of the effects of ageing on biometric identity verification," *International Journal of Biometrics*, vol. 2, no. 1, pp. 34–52, 2009.

[18] A. Lanitis and N. Tsapatsoulis, "Quantitative evaluation of the effects of ageing on biometric templates," *Computer Vision, IET*, vol. 5, no. 6, pp. 338–347, 2011.

[19] G. Guo and G. Mu, "Human age estimation: What is the influence across race and gender?" in *Computer Vision and Pattern Recognition Workshops (CVPRW), 2010 IEEE Computer Society Conference on.* IEEE, 2010, pp. 71–78.

[20] N. Ramanathan and R. Chellappa, "Modeling age progression in young faces," in *Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on*, vol. 1. IEEE, 2006, pp. 387–394.

[21] A. Tsymbal, "The problem of concept drift: definitions and related work," *Computer Science Department, Trinity College Dublin*, vol. 106, p. 2, 2004.

[22] N. Dayan, "Skin ageing handbook: Market perspectives, pharmacology, formulation, and evaluation techniques," 2008.

[23] S. R. Coleman and R. Grover, "The anatomy of the ageing face: volume loss and changes in 3-dimensional topography," *Aesthetic surgery journal*, vol. 26, no. 1 Supplement, pp. S4–S9, 2006.

[24] M. S. Zimbler, M. S. Kokoska, and J. R. Thomas, "Anatomy and pathophysiology of facial ageing." *Facial plastic surgery clinics of North America*, vol. 9, no. 2, pp. 179–87, 2001.

[25] M. A. Taister, S. D. Holliday, and H. Borrman, "Comments on facial ageing in law enforcement investigation," *Forensic science communications*, vol. 2, no. 2, 2000.

[26] R. Pandel, B. Poljšak, A. Godic, and R. Dahmane, "Skin photoaging and the role of antioxidants in its prevention," *ISRN dermatology*, vol. 2013, 2013.

[27] T. Overfield, *Biological variation in health and illness: Race, age, and sex differences.* CRC Press, 1995.

[28] Y. M. Lui, D. Bolme, P. J. Phillips, J. R. Beveridge, and B. A. Draper, "Preliminary studies on the good, the bad, and the ugly face recognition challenge

problem," in *Computer Vision and Pattern Recognition Workshops (CVPRW), 2012 IEEE Computer Society Conference on.* IEEE, 2012, pp. 9–16.

[29] A. S. Georghiades, P. N. Belhumeur, and D. J. Kriegman, "From few to many: Generative models for recognition under variable pose and illumination," in *Automatic Face and Gesture Recognition, 2000. Proceedings. Fourth IEEE International Conference on.* IEEE, 2000, pp. 277–284.

[30] Y. Adini, Y. Moses, and S. Ullman, "Face recognition: The problem of compensating for changes in illumination direction," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 19, no. 7, pp. 721–732, 1997.

[31] R. Chellappa, C. L. Wilson, and S. Sirohey, "Human and machine recognition of faces: A survey," *Proceedings of the IEEE*, vol. 83, no. 5, pp. 705–741, 1995.

[32] V. Govindaraju, S. N. Srihari, and D. B. Sher, "A computational model for face location," in *Computer Vision, 1990. Proceedings, Third International Conference on.* IEEE, 1990, pp. 718–721.

[33] C. J. Wu and J. S. Huang, "Human face profile recognition by computer," *Pattern recognition*, vol. 23, no. 3, pp. 255–259, 1990.

[34] G. G. Gordon, "Face recognition based on depth maps and surface curvature," in *San Diego,'91, San Diego, CA.* International Society for Optics and Photonics, 1991, pp. 234–247.

[35] E. Patterson, A. Sethuram, M. Albert, and K. Ricanek, "Comparison of synthetic face ageing to age progression by forensic sketch artist," in *IASTED International Conference on Visualization, Imaging, and Image Processing, Palma de Mallorca, Spain*, 2007.

[36] K. Ricanek Jr, G. Mahalingam, A. M. Albert, and R. W. V. Bruegge, "Human face ageing: a perspective analysis from anthropometry and biometrics," *Age factors in biometric processing(IET, UK, 2013)*, pp. 93–116, 2013.

[37] A. Lanitis, N. Tsapatsoulis, K. Soteriou, D. Kuwahara, and S. Morishima, "Fg2015 age progression evaluation," in *Automatic Face and Gesture Recognition (FG), 2015 11th IEEE International Conference and Workshops on.* IEEE, 2015, pp. 1–6.

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

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

[40] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, vol. 1. IEEE, 2005, pp. 886–893.

[41] T. H. N. Le, K. Seshadri, K. Luu, and M. Savvides, "Facial ageing and asymmetry decomposition based approaches to identification of twins," *Pattern Recognition*, vol. 48, no. 12, pp. 3843–3856, 2015.

[42] K. Seshadri and M. Savvides, "Robust modified active shape model for automatic facial landmark annotation of frontal faces," in *Biometrics: Theory, Applications, and Systems, 2009. BTAS'09. IEEE 3rd International Conference on.* IEEE, 2009, pp. 1–8.

[43] H. Wang, S. Z. Li, Y. Wang, and J. Zhang, "Self quotient image for face recognition," in *Image Processing, 2004. ICIP'04. 2004 International Conference on*, vol. 2. IEEE, 2004, pp. 1397–1400.

[44] A. Lanitis, C. J. Taylor, and T. F. Cootes, "Toward automatic simulation of ageing effects on face images," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 24, no. 4, pp. 442–455, 2002.

[45] D. M. Burt and D. I. Perrett, "Perception of age in adult caucasian male faces: Computer graphic manipulation of shape and colour information," *Proceedings*

of the Royal Society of London B: Biological Sciences, vol. 259, no. 1355, pp. 137–143, 1995.

[46] M. R. Gandhi, "A method for automatic synthesis of aged human facial images," Ph.D. dissertation, McGill University, Montreal, Canada, 2004.

[47] B. Tiddeman, M. Burt, and D. Perrett, "Prototyping and transforming facial textures for perception research," *Computer Graphics and Applications, IEEE*, vol. 21, no. 5, pp. 42–50, 2001.

[48] A. Lanitis, C. J. Taylor, and T. F. Cootes, "Automatic interpretation and coding of face images using flexible models," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 19, no. 7, pp. 743–756, 1997.

[49] E. J. Stollnitz, T. D. DeRose, and D. H. Salesin, "Wavelets for computer graphics: A primer, part 1 & 2," *IEEE Computer Graphics and Applications*, vol. 15, no. 3, 1996.

[50] K. Ricanek Jr and E. Boone, "The effect of normal adult ageing on standard pca face recognition accuracy rates," in *Neural Networks, 2005. IJCNN'05. Proceedings. 2005 IEEE International Joint Conference on*, vol. 4. IEEE, 2005, pp. 2018–2023.

[51] M. Turk and A. Pentland, "Eigenfaces for recognition," *Journal of cognitive neuroscience*, vol. 3, no. 1, pp. 71–86, 1991.

[52] D. A. Rowland and D. I. Perrett, "Manipulating facial appearance through shape and color," *Computer Graphics and Applications, IEEE*, vol. 15, no. 5, pp. 70–76, 1995.

[53] I. Kemelmacher-Shlizerman, S. Suwajanakorn, and S. M. Seitz, "Illumination-aware age progression," in *Computer Vision and Pattern Recognition (CVPR), 2014 IEEE Conference on*. IEEE, 2014, pp. 3334–3341.

[54] A. Sethuram, K. Ricanek, and E. Patterson, "A hierarchical approach to facial ageing," in *Computer Vision and Pattern Recognition Workshops (CVPRW), 2010 IEEE Computer Society Conference on*. IEEE, 2010, pp. 100–107.

[55] A. Sethuram, E. Patterson, K. Ricanek, and A. Rawls, "Improvements and performance evaluation concerning synthetic age progression and face recognition affected by adult ageing," in *Advances in Biometrics*. Springer, 2009, pp. 62–71.

[56] R. Beveridge, D. Bolme, M. Teixeira, and B. Draper, "The csu face identification evaluation system users guide: version 5.0," *Computer Science Department, Colorado State University*, vol. 2, no. 3, 2003.

[57] N. Ramanathan and R. Chellappa, "Modeling shape and textural variations in ageing faces," in *Automatic Face & Gesture Recognition, 2008. FG'08. 8th IEEE International Conference on*. IEEE, 2008, pp. 1–8.

[58] "Ageing booth. version 2.6," 2014, retrieved from: http://www.piviandco.com/apps/agingbooth/.

[59] Y. Wang, Z. Zhang, W. Li, and F. Jiang, "Combining tensor space analysis and active appearance models for ageing effect simulation on face images," *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, vol. 42, no. 4, pp. 1107–1118, 2012.

[60] A. Maronidis and A. Lanitis, "Facial age simulation using age-specific 3d models and recursive pca." in *VISAPP (1)*, 2013, pp. 663–668.

[61] G. Panis, A. Lanitis, N. Tsapatsoulis, and T. F. Cootes, "Overview of research on facial ageing using the fg-net ageing database," *IET Biometrics*, 2015.

[62] C. M. Scandrett, C. J. Solomon, and S. J. Gibson, "A person-specific, rigorous ageing model of the human face," *Pattern Recognition Letters*, vol. 27, no. 15, pp. 1776–1787, 2006.

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

[64] J. Suo, X. Chen, S. Shan, W. Gao, and Q. Dai, "A concatenational graph evolution ageing model," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 34, no. 11, pp. 2083–2096, 2012.

[65] K. Luu, T. D. Bui, C. Y. Suen, and K. Ricanek, “Spectral regression based age determination,” in *Computer Vision and Pattern Recognition Workshops (CVPRW), 2010 IEEE Computer Society Conference on.* IEEE, 2010, pp. 103–107.

[66] M.-H. Tsai, Y.-K. Liao, and I.-C. Lin, “Human face ageing with guided prediction and detail synthesis,” *Multimedia tools and applications*, vol. 72, no. 1, pp. 801–824, 2014.

[67] B. M. Barer, “Men and women ageing differently.” *The International Journal of Ageing & Human Development*, 1994.

[68] I. Prakash, “Women & ageing.” *The Indian journal of medical research*, vol. 106, pp. 396–408, 1997.

[69] P. L. Leong, “Ageing changes in the male face,” *Facial plastic surgery clinics of North America*, vol. 16, no. 3, pp. 277–279, 2008.

[70] E. C. Paes, H. J. Teepen, W. A. Koop, and M. Kon, “Perioral wrinkles: histologic differences between men and women,” *Aesthetic Surgery Journal*, vol. 29, no. 6, pp. 467–472, 2009.

[71] D. A. Gunn, H. Rexbye, C. Griffiths, P. G. Murray, A. Fereday, S. D. Catt, C. C. Tomlin, B. H. Strongitharm, D. I. Perrett, M. Catt *et al.*, “Why some women look young for their age,” *PloS one*, vol. 4, no. 12, p. e8021, 2009.

[72] H. Rexbye, I. Petersen, M. Johansens, L. Klitkou, B. Jeune, and K. Christensen, “Influence of environmental factors on facial ageing,” *Age and Ageing*, vol. 35, no. 2, pp. 110–115, 2006.

[73] J. Uitto, “Understanding premature skin ageing,” *New England Journal of Medicine*, vol. 337, no. 20, pp. 1463–1465, 1997.

[74] R. B. Shaw Jr, E. B. Katzel, P. F. Koltz, D. M. Kahn, J. A. Girotto, and H. N. Langstein, “Ageing of the mandible and its aesthetic implications,” *Plastic and reconstructive surgery*, vol. 125, no. 1, pp. 332–342, 2010.

[75] M. Šitum, M. Buljan, V. Čavka, V. Bulat, I. Krolo, and L. Lugović Mihić, "Skin changes in the elderly people–how strong is the influence of the uv radiation on skin ageing?" *Collegium antropologicum*, vol. 34, no. 2, pp. 9–13, 2010.

[76] R. Neave, "Age changes to the face in adulthood," *Craniofacial Identification in Forensic Medicine*, pp. 215–231, 1998.

[77] G. Hillebrand, M. Levine, and K. Miyamoto, "The age-dependent changes in skin condition in african americans, asian indians, caucasians, east asians and latinos," *IFSCC Mag*, vol. 4, no. 4, p. 2591266, 2001.

[78] R. Singh and H. Om, "An overview of face recognition in an unconstrained environment," in *2013 IEEE Second International Conference on Image Information Processing (ICIIP-2013)*, 2013.

[79] N. Isse, "Silhouette sutures for treatment of facial ageing: facial rejuvenation, remodeling, and facial tissue support," *Clinics in plastic surgery*, vol. 35, no. 4, pp. 481–486, 2008.

[80] N. S. Sadick, C. Karcher, and L. Palmisano, "Cosmetic dermatology of the ageing face," *Clinics in dermatology*, vol. 27, no. 3, pp. S3–S12, 2009.

[81] S. Victor, "You under the sun," *Victoria*, vol. 17, pp. 46–48, 2003.

[82] A. K. Jain, R. P. Duin, and J. Mao, "Statistical pattern recognition: A review," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 22, no. 1, pp. 4–37, 2000.

[83] R. Brunelli and T. Poggio, "Face recognition: Features versus templates," *IEEE Transactions on Pattern Analysis & Machine Intelligence*, no. 10, pp. 1042–1052, 1993.

[84] B. Heisele, P. Ho, and T. Poggio, "Face recognition with support vector machines: Global versus component-based approach," in *Computer Vision, 2001. ICCV 2001. Proceedings. Eighth IEEE International Conference on*, vol. 2. IEEE, 2001, pp. 688–694.

[85] G. Guo, S. Z. Li, and K. Chan, "Face recognition by support vector machines," in *Automatic Face and Gesture Recognition, 2000. Proceedings. Fourth IEEE International Conference on.* IEEE, 2000, pp. 196–201.

[86] K. Jonsson, J. Matas, J. Kittler, and Y. Li, "Learning support vectors for face verification and recognition," in *Automatic Face and Gesture Recognition, 2000. Proceedings. Fourth IEEE International Conference on.* IEEE, 2000, pp. 208–213.

[87] M. Lades, J. C. Vorbrüggen, J. Buhmann, J. Lange, C. V. d Malsburg, R. Wurtz, and W. Konen, "Distortion invariant object recognition in the dynamic link architecture," *Computers, IEEE Transactions on*, vol. 42, no. 3, pp. 300–311, 1993.

[88] L. D. Harmon, A. B. Lesk *et al.*, "Identification of human faces," *Proceedings of the IEEE*, vol. 59, no. 5, pp. 748–760, 1971.

[89] A. Samal and P. A. Iyengar, "Automatic recognition and analysis of human faces and facial expressions: A survey," *Pattern recognition*, vol. 25, no. 1, pp. 65–77, 1992.

[90] I. J. Cox, J. Ghosn, and P. N. Yianilos, "Feature-based face recognition using mixture-distance," in *Computer Vision and Pattern Recognition, 1996. Proceedings CVPR'96, 1996 IEEE Computer Society Conference on.* IEEE, 1996, pp. 209–216.

[91] W. Zhao, R. Chellappa, P. J. Phillips, and A. Rosenfeld, "Face recognition: A literature survey," *ACM computing surveys (CSUR)*, vol. 35, no. 4, pp. 399–458, 2003.

[92] B. Manjunath, R. Chellappa, and C. von der Malsburg, "A feature based approach to face recognition," in *Computer Vision and Pattern Recognition, 1992. Proceedings CVPR'92., 1992 IEEE Computer Society Conference on.* IEEE, 1992, pp. 373–378.

[93] R. J. Baron, "Mechanisms of human facial recognition," *International Journal of Man-Machine Studies*, vol. 15, no. 2, pp. 137–178, 1981.

[94] L. Sirovich and M. Kirby, "Low-dimensional procedure for the characterization of human faces," *JOSA A*, vol. 4, no. 3, pp. 519–524, 1987.

[95] K. Etemad and R. Chellappa, "Discriminant analysis for recognition of human face images," *JOSA A*, vol. 14, no. 8, pp. 1724–1733, 1997.

[96] A. M. Martínez and A. C. Kak, "Pca versus lda," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 23, no. 2, pp. 228–233, 2001.

[97] K. Person, "On lines and planes of closest fit to system of points in space. philiosophical magazine, 2, 559-572," 1901.

[98] S. Watanabe, "Karhunen-loeve expansion and factor analysis, theoretical remarks and applications," in *Proc. 4th Prague Conf. Inform. Theory*, 1965.

[99] M. Kirby and L. Sirovich, "Application of the karhunen-loeve procedure for the characterization of human faces," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 12, no. 1, pp. 103–108, 1990.

[100] K. Chang, K. W. Bowyer, S. Sarkar, and B. Victor, "Comparison and combination of ear and face images in appearance-based biometrics," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 25, no. 9, pp. 1160–1165, 2003.

[101] K.-K. Sung, "Learning and example selection for object and pattern detection," 1996.

[102] F. Soulie, "Connectionist methods for human face processing," in *Nato Asi Conference on Faces*, 1997.

[103] H. A. Rowley, S. Baluja, and T. Kanade, "Neural network-based face detection," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 20, no. 1, pp. 23–38, 1998.

[104] F. Raphaël, B. Olivier, and C. Daniel, "A constrained generative model applied to face detection," *Neural Processing Letters*, vol. 5, no. 2, pp. 11–19, 1997.

[105] C.-H. Chan, J. Kittler, and K. Messer, *Multi-scale local binary pattern histograms for face recognition.* Springer, 2007.

[106] T. Ahonen, E. Rahtu, V. Ojansivu, and J. Heikkilä, "Recognition of blurred faces using local phase quantization," in *Pattern Recognition, 2008. ICPR 2008. 19th International Conference on.* IEEE, 2008, pp. 1–4.

[107] L. Wiskott, J.-M. Fellous, N. Kuiger, and C. Von Der Malsburg, "Face recognition by elastic bunch graph matching," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 19, no. 7, pp. 775–779, 1997.

[108] D. R. Kisku, A. Rattani, E. Grosso, and M. Tistarelli, "Face identification by sift-based complete graph topology," in *Automatic Identification Advanced Technologies, 2007 IEEE Workshop on.* IEEE, 2007, pp. 63–68.

[109] N. Ramanathan, R. Chellappa, S. Biswas *et al.*, "Age progression in human faces: A survey," *Journal of Visual Languages and Computing*, vol. 15, pp. 3349–3361, 2009.

[110] A. K. Jain, S. C. Dass, and K. Nandakumar, "Soft biometric traits for personal recognition systems," in *Biometric Authentication.* Springer, 2004, pp. 731–738.

[111] E. Patterson, K. Ricanek, M. Albert, and E. Boone, "Automatic representation of adult ageing in facial images," in *Proc. IASTED Intl Conf. Visualization, Imaging, and Image Processing*, 2006, pp. 171–176.

[112] Y. Fu, Y. Xu, and T. S. Huang, "Estimating human age by manifold analysis of face pictures and regression on ageing features," in *Multimedia and Expo, 2007 IEEE International Conference on.* IEEE, 2007, pp. 1383–1386.

[113] G. Guo, Y. Fu, T. S. Huang, and C. R. Dyer, "Locally adjusted robust regression for human age estimation," 2008.

[114] B. Ni, Z. Song, and S. Yan, "Web image mining towards universal age estimator," in *Proceedings of the 17th ACM international conference on Multimedia.* ACM, 2009, pp. 85–94.

[115] G. Panis and A. Lanitis, "An overview of research activities in facial age estimation using the fg-net ageing database," in *Computer Vision-ECCV 2014 Workshops.* Springer, 2014, pp. 737–750.

[116] Face, G. R. W. Group *et al.*, "The fg-net ageing database," 2000, [2011-06-28]. Retrieved From: http://www, fgnet. rsunit, cow.

[117] A. Gallagher and T. Chen, "Understanding images of groups of people," in *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on.* IEEE, 2009, pp. 256–263.

[118] A. C. Gallagher and T. Chen, "Clothing cosegmentation for recognizing people," in *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on.* IEEE, 2008, pp. 1–8.

[119] K. Ueki, T. Hayashida, and T. Kobayashi, "Subspace-based age-group classification using facial images under various lighting conditions," in *Automatic Face and Gesture Recognition, 2006. FGR 2006. 7th International Conference on.* IEEE, 2006, pp. 6–pp.

[120] J. Suo, F. Min, S. Zhu, S. Shan, and X. Chen, "A multi-resolution dynamic model for face ageing simulation," in *Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on.* IEEE, 2007, pp. 1–8.

[121] S. J. Foundation, "Human and object interaction processing (hoip) face database," 2010.

[122] Y. Fu and N. Zheng, "M-face: An appearance-based photorealistic model for multiple facial attributes rendering," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 16, no. 7, pp. 830–842, 2006.

[123] Y. Fu and T. S. Huang, "Human age estimation with regression on discriminative ageing manifold," *Multimedia, IEEE Transactions on*, vol. 10, no. 4, pp. 578–584, 2008.

[124] K. Scherbaum, M. Sunkel, H.-P. Seidel, and V. Blanz, "Prediction of individual non-linear ageing trajectories of faces," in *Computer Graphics Forum*, vol. 26, no. 3. Wiley Online Library, 2007, pp. 285–294.

[125] V. Blanz and T. Vetter, "A morphable model for the synthesis of 3d faces," in *Proceedings of the 26th annual conference on Computer graphics and interactive techniques.* ACM Press/Addison-Wesley Publishing Co., 1999, pp. 187–194.

[126] P. J. Phillips, P. J. Flynn, T. Scruggs, K. W. Bowyer, J. Chang, K. Hoffman, J. Marques, J. Min, and W. Worek, "Overview of the face recognition grand challenge," in *Computer vision and pattern recognition, 2005. CVPR 2005. IEEE computer society conference on*, vol. 1. IEEE, 2005, pp. 947–954.

[127] T. Acharya and A. K. Ray, *Image processing: principles and applications.* John Wiley & Sons, 2005.

[128] P. Griffin, "Understanding the face image format standards," in *ANSI/NIST Workshop, Gaithersburg, MD*, 2005.

[129] T. Aach, A. Kaup, and R. Mester, "On texture analysis: Local energy transforms versus quadrature filters," *Signal Processing*, vol. 45, no. 2, pp. 173–181, 1995.

[130] M. R. Jenkin and A. D. Jepson, "Recovering local surface structure through local phase difference measurements," *CVGIP: Image understanding*, vol. 59, no. 1, pp. 72–93, 1994.

[131] A. Maki, L. Bretzner, and J.-O. Eklundh, "Local fourier phase and disparity estimates: An analytical study," in *Computer Analysis of Images and Patterns.* Springer, 1995, pp. 868–873.

[132] D. J. Fleet and A. D. Jepson, "Stability of phase information," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 15, no. 12, pp. 1253–1268, 1993.

[133] D. J. Fleet, A. D. Jepson, and M. R. Jenkin, "Phase-based disparity measurement," *CVGIP: Image understanding*, vol. 53, no. 2, pp. 198–210, 1991.

[134] C. Liu, "Capitalize on dimensionality increasing techniques for improving face recognition grand challenge performance," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 28, no. 5, pp. 725–737, 2006.

[135] J. G. Daugman, "Complete discrete 2-d gabor transforms by neural networks for image analysis and compression," *Acoustics, Speech and Signal Processing, IEEE Transactions on*, vol. 36, no. 7, pp. 1169–1179, 1988.

[136] R. Hutchinson and W. Welsh, "Comparison of neural networks and conventional techniques for feature location in facial images," in *Artificial Neural Networks, 1989., First IEE International Conference on (Conf. Publ. No. 313).* IET, 1989, pp. 201–205.

[137] T. Ojala, M. Pietikäinen, and D. Harwood, "A comparative study of texture measures with classification based on featured distributions," *Pattern recognition*, vol. 29, no. 1, pp. 51–59, 1996.

[138] L. Wang and D.-C. He, "Texture classification using texture spectrum," *Pattern Recognition*, vol. 23, no. 8, pp. 905–910, 1990.

[139] T. Ojala, M. Pietikäinen, and T. Mäenpää, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 24, no. 7, pp. 971–987, 2002.

[140] T. Mäenpää and M. Pietikäinen, "Texture analysis with local binary patterns," *Handbook of Pattern Recognition and Computer Vision*, vol. 3, pp. 197–216, 2005.

[141] T. Ahonen, A. Hadid, and M. Pietikäinen, "Face recognition with local binary patterns," in *Computer vision-eccv 2004*. Springer, 2004, pp. 469–481.

[142] S. Pawan, B. Balas, Y. Ostrovsky, and R. Russell, "Face recognition by humans: 20 results all computer vision researchers should know about," *Department of Brain and Cognitive Sciences Massachusetts Institute of Technology Cambridge, MA*, vol. 2139, 2005.

[143] M. S. Keil, "I look in your eyes, honey: Internal face features induce spatial frequency preference for human face processing," *PLoS Comput. Biol*, vol. 5, no. 3, pp. 1–13, 2009.

[144] J. Shlens, "A tutorial on principal component analysis," *arXiv preprint arXiv:1404.1100*, 2014.

[145] W. Yambor, B. Draper, and R. Beveridge, "Analyzing pca-based face recognition algorithms: Eigenvector selection and distance measures," in *2nd Workshop on*

Empirical Evaluation in Computer Vision, Dublin, Ireland. World Scientific, 2000.

[146] S. Balakrishnama and A. Ganapathiraju, "Linear discriminant analysis-a brief tutorial," *Institute for Signal and information Processing*, 1998.

[147] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern classification.* John Wiley & Sons, 2012.

[148] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, "Eigenfaces vs. fisherfaces: Recognition using class specific linear projection," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 19, no. 7, pp. 711–720, 1997.

[149] B. R. Corner, R. M. Narayanan, and S. E. Reichenbach, "Principal component analysis of remote sensing imagery: effects of additive and multiplicative noise," in *SPIE's International Symposium on Optical Science, Engineering, and Instrumentation.* International Society for Optics and Photonics, 1999, pp. 183–191.

[150] B. Schölkopf, A. Smola, and K.-R. Müller, "Nonlinear component analysis as a kernel eigenvalue problem," *Neural computation*, vol. 10, no. 5, pp. 1299–1319, 1998.

[151] M. Kubat, "Neural networks: a comprehensive foundation by simon haykin, macmillan, 1994, isbn 0-02-352781-7." 1999.

[152] S. Haykin and N. Network, "A comprehensive foundation," *Neural Networks*, vol. 2, no. 2004, 2004.

[153] J. Shawe-Taylor and N. Cristianini, *Kernel methods for pattern analysis.* Cambridge university press, 2004.

[154] C. Liu, "Gabor-based kernel pca with fractional power polynomial models for face recognition," *IEEE transactions on pattern analysis and machine intelligence*, vol. 26, no. 5, pp. 572–581, 2004.

[155] P.-E. Danielsson, "Euclidean distance mapping," *Computer Graphics and image processing*, vol. 14, no. 3, pp. 227–248, 1980.

[156] P. Das, "Lattice of octagonal distances in digital geometry," *Pattern Recognition Letters*, vol. 11, no. 10, pp. 663–667, 1990.

[157] H. V. Nguyen and L. Bai, "Cosine similarity metric learning for face verification," in *Computer Vision–ACCV 2010*. Springer, 2010, pp. 709–720.

[158] J. Cook, C. McCool, V. Chandran, and S. Sridharan, "Combined 2d/3d face recognition using log-gabor templates," in *Video and Signal Based Surveillance, 2006. AVSS'06. IEEE International Conference on*. IEEE, 2006, pp. 83–83.

[159] M. Pietikäinen, A. Hadid, G. Zhao, and T. Ahonen, "Local binary patterns for still images," in *Computer Vision Using Local Binary Patterns*. Springer, 2011, pp. 13–47.

[160] C. Liu and H. Wechsler, "A gabor feature classifier for face recognition," in *Computer Vision, 2001. ICCV 2001. Proceedings. Eighth IEEE International Conference on*, vol. 2. IEEE, 2001, pp. 270–275.

[161] L. Chengjun and W. Harry, "Independent component analysis of gabor features for face recognition," *IEEE transactions on Neural Networks*, vol. 14, no. 4, pp. 919–928, 2003.

[162] W. Zhang, S. Shan, W. Gao, X. Chen, and H. Zhang, "Local gabor binary pattern histogram sequence (lgbphs): a novel non-statistical model for face representation and recognition," in *Tenth IEEE International Conference on Computer Vision (ICCV'05) Volume 1*, vol. 1. IEEE, 2005, pp. 786–791.

[163] B. Xia, H. Sun, and B.-L. Lu, "Multi-view gender classification based on local gabor binary mapping pattern and support vector machines," in *2008 IEEE International Joint Conference on Neural Networks (IEEE World Congress on Computational Intelligence)*. IEEE, 2008, pp. 3388–3395.

[164] C. MageshKumar, G. Ragul, and R. Thiyagarajan, "Gabor featured statistical modeling in face recognition with chaotic database," vol. 3, January 2013.

[165] Y. Sun, J.-M. Zhang, L.-M. Wang, Y.-Z. Zhan, and S.-L. Song, "A novel method of recognizing ageing face based on ehmm," in *Machine Learning and Cybernetics, 2005. Proceedings of 2005 International Conference on*, vol. 8. IEEE, 2005, pp. 4599–4604.

[166] M. Erbilek and M. Fairhurst, "A methodological framework for investigating age factors on the performance of biometric systems," in *Proceedings of the on Multimedia and security.* ACM, 2012, pp. 115–122.

[167] R. Singh, M. Vatsa, A. Noore, and S. K. Singh, "Age transformation for improving face recognition performance," in *Pattern Recognition and Machine Intelligence.* Springer, 2007, pp. 576–583.

[168] W. H. Ho, P. Watters, and D. Verity, "Are younger people more difficult to identify or just a peer-to-peer effect," in *Computer Analysis of Images and Patterns.* Springer, 2007, pp. 351–359.

[169] Y. M. Lui, D. Bolme, B. A. Draper, J. R. Beveridge, G. Givens, and P. J. Phillips, "A meta-analysis of face recognition covariates," in *IEEE 3rd International Conference on Biometrics: Theory, Applications, and Systems, Washington, DC*, 2009.

[170] C. M. Bishop, *Neural networks for pattern recognition.* Oxford university press, 1995.

[171] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning internal representations by error propagation," DTIC Document, Tech. Rep., 1985.

[172] I. Yilmaz and G. Yuksek, "Prediction of the strength and elasticity modulus of gypsum using multiple regression, ann, and anfis models," *International Journal of Rock Mechanics and Mining Sciences*, vol. 46, no. 4, pp. 803–810, 2009.

[173] O. N. Al-allaf, A. A. Tamimi, and M. A. Alia, "Face recognition system based on different artificial neural networks models and training algorithms," *International Journal of Advanced Computer Science and Applications*, vol. 4, no. 6, 2013.

[174] S. Godara and R. Gupta, "Neural networks for iris recognition: Comparisons between lvq and cascade forward back propagation neural network models, architectures and algorithm," *Neural Networks*, vol. 3, no. 1, 2013.

[175] O. N. A. Al-allaf, A. A. Tamimi, and S. A. AbdAlKader, "Artificial neural networks for iris recognition system: Comparisons between different models, architectures and algorithms 1," 2012.