Age Invariant Face Recognition via Aging Modelling

By

Fahad Bashir Alvi

Primary Supervisor: Associate Professor Russel Pears

Secondary Supervisor : Professor Nikola Kasabov

SUBMITTED FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY
AT
SCHOOL OF ENGINEERING, COMPUTER AND MATHEMATICAL SCIENCES
AUCKLAND UNIVERSITY OF TECHNOLOGY
AUCKLAND, NEW ZEALAND
JUNE 2017

To my parents

List of Tables

List of Figures

Attestation of Authorship

I hereby declare that this submission is my own work and that, to the best of my knowledge and belief, it contains no material previously published or written by another person (except where explicitly defined in the acknowledgements), nor material which to a substantial extent has been submitted for the award of any other degree or diploma of a university or other institution of higher learning.

Fahad Bashir Alvi
AUT, New Zealand.

Abstract

Aging is a complex problem because at different age points different changes occur in the human face. From childhood to teenage the changes are mostly related to craniofacial growth. At maturity the changes are mostly related to the skin color changes and the face skin starts becoming slack and less smooth. So aging is a mixture of all these components. Moreover, aging is a slow, irreversible and a process that is unique to every human being. Many factors affect the aging process. For example every person has different genes, blood group, and life style and belongs to a particular ethnic group. In order to resolve all these issues we need modelling method that should capture all changes throughout the aging process. So we deal shape and texture separately and finally combine them to develop a synergy between the two. The main objective of thesis is to develop novel aging growth models by using spatiotemporal modeling. It aims to develop and enhance our knowledge of craniofacial growth of human faces and capturing it to form models to understand the changes occurring due to aging process. We investigate five different types of models built on different levels of data granularity and with different perspective. At the global level a model is built on training data that encompasses the entire set of available individuals, whereas at the local level data from homogeneous sub-populations is used and finally at the individual level a personalized model is built for each individual. We used anthropometric features for extracting the shape information of face and for texture information we extract features based on the edges on the texture of the human face. Finally, we make integrated model in order to utilize synergy between three models and arrive at an optimum solution. Our Integrated model is inspired by the Integrated

Multi Model Frame Work(IMMF) proposed by Widuputra which exploits synergy that exists between models at the global, local and personalized levels. Our integrated model is based on Adaptive Linear Neuron or later Adaptive Linear Element (Adaline) by calculating weights of each model. A concept of deaging was introduced to align and make our features commensurate with different ages. We built a novel face recognition system based on aging models. It is a two stage process. At first stage we built the model and in second stage we try to identify the probe image into the subspace of the gallery. In the second stage real problem is finding the correct subspace. We have used Naive Bayesian method and the probe image subspace for finding the most probable subspace. The probe image is then searched only in the reduced search space of already determined subspace instead of the whole gallery, thus resulting in reduction of cost of computation time. We built two different sets of models, one for the shape and the other for texture. It was observed that we need to resolve the conflict if our models point to different persons. We enunciated a method to resolve the conflict between two models by building composite models in which we combined the texture and shape. Our composite model is developed by using Decision Tree approach. We have used well known standard databases related to aging, namely, FG-NET and MORPH. Our experimentation showed good results as well as demonstrated utility of our aging models. We also know that aging is primarily a spatiotemporal process. So, tools for spatiotemporal and temporal data (STTD) were also used. The Knowledge Engineering and Discovery Research Institute (KEDRI) has developed methods and tools to deal with STTD, the latest being the NeuCube. We built NeuCube (ST) aging model for age group classification and gender recognition thus demonstrating utilization of a robust tool viz NeuCube (ST).

Acknowledgments

All praises to the Allah/God Almighty who induced the man with intelligence, knowledge and wisdom. It is He who gave me ability, perseverance and determination to complete this work successfully.

It gives me great pleasure in offering my profound gratitude to my primary supervisor Associate Professor. Russel Pears, for his inspiring guidance, continuous encouragement and valuable comments for reviewing my manuscript. Specially, weekly meeting with Prof. Russel helped me a lot in learning skills and keeping me on track. I am fortunate to have had Prof. Russel as my supervisor.

I would like to thank my secondary supervisor Professor. Nikola Kasabov for his motivation and encouragement throughout my PhD studies. Specially in weekly KEDRI meetings, I had knowledgeable interaction with him and his vision opened new vistas of knowledge for me. I got opportunity to learn a lot from him.

I would like to proffer special thanks to Joyce D Mello, for her encouraging words and guidance throughout my time in KEDRI. I also wish to thank other members of KEDRI at the Auckland University of Technology, who provided me with good research environment.

I offer my thanks to all the people and organizations who collected and provided FG-NET and MORPH databases for research.

Last but not the least my special thanks to my parents, my wife and my daughters Maham, Maheen and Fizza, my brother, sisters and friends for their special prayers and encouragement. Their un-conditional love and undoubted belief in me have sustained me throughout this work.

Auckland, New Zealand
June 2017

Fahad Bashir Alvi

Chapter 1

Introduction

1.1 Back ground

The history of biometrics can be said to begin with the development of the Bertillion system. A French police officer used biometrics information to identify criminals (Rhodes, 1956). Generally, six biometric attributes, viz face, finger, hand, voice , eye, ear signatures are used for biometrics(Heitmeyer, 2000). It is noted that face based recognition systems are most convenient because it is natural, nonintrusive and easy to use (A. Jain, Flynn, & Ross, 2007; A. K. Jain & Li, 2011). The recognition system based on the human face has been around for many years.

Face image recognition is an important and significant part of the domain of biometrics reseach(A. K. Jain, Nandakumar, & Ross, 2016; Cootes, Edwards, Taylor, et al., 1999; Gong, Li, Tao, Liu, & Li, 2015; Li, Park, & Jain, 2011; Ramanathan & Chellappa, 2006; Park, Tong, & Jain, 2010). Face recognition has two main areas of research, namely face verification and face identification. In face recognition system, face images from a number of subjects are enrolled into the system as gallery data, and the face image of a test subject (probe image) is matched to the gallery data using a one-to-one or one-to- many scheme. The one-to-one and one-to-many matchings

are called verification and identification, respectively. Figure 1.1 shows the difference between verification and identification system. Face recognition can be performed in open set or closed set scenarios. Closed set recognition tests always have the probe subject enrolled in the gallery data, but the open set recognition consider the possibility that the probe subject is not enrolled in the gallery. Therefore, a threshold value (on match score) is typically used in retrieving candidate matches in open set recognition tests (A. Jain et al., 2007; A. K. Jain & Li, 2011).

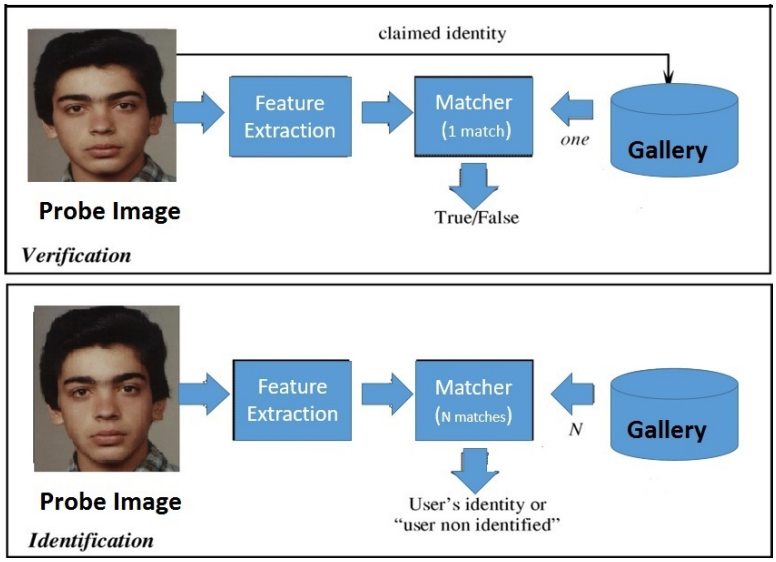


Figure 1.1: Difference between verification and identification

Face recognition is an extensively researched topic.The first automated system utilizing face recognition was developed in 1977 (Kanade, 1973). Although there has been significant improvements in face recognition performance during the past decade, it is still below an acceptable level. Many different algorithms for face recognition exist, related to different fields like machine learning, pattern recognition and computer vision. The first algorithm was Principal Component Analysis (PCA) which describes

the faces as a linear combination of eigenfaces(Sirovich & Kirby, 1987). Further, Linear Discriminant Analysis (LDA) was introduced which is based on fisherfaces introduced as improvement of PCA approach(Belhumeur, Hespanha, & Kriegman, 1997). Independent Component Analysis (ICA) followed, which make use of higher-order statistics to improve the system(Déniz, Castrillon, & Hernández, 2003). Many variations of the PCA, LDA and ICA algorithms have been made with varying results (Delac, Grgic, & Grgic, 2005a, 2005b; Nefian & Hayes III, 1998). Researchers studied different techniques and methods in order to increase the accuracy because existing techniques were not suitable for combating confounding effects arising from illumuniation, expression, occlusion, pose and aging. These techniques include such as Local Binary Patterns (LBP) (Zhang, Huang, Li, Wang, & Wu, 2004), Gabbor wavelets and Elastic Bunch Graph Matching (EBGM)(Bolme, 2003) and Active Appearance Model (AAM) etc. Despite all the research and many different algorithms, new methods that deal with variations in expression, pose and aging are still required. Different evaluations such as FERET (Phillips, Moon, Rizvi, & Rauss, 2000; Phillips, Moon, Rauss, & Rizvi, 1997), FRVT (Phillips et al., 2003) and FRVT(Phillips et al., 2010) have proved that aging, illumination and pose are still plaguing the performance of state of art algorithms(Champod, 2015).

This thesis will research the issue of age modelling, and in particular the age invariant face recognition problem. Novel methods for age invariant face recognition are presented and an in-depth expeirmental study of these methods on well established benchmark datasets are presented.

1.2 Challenges in Face Recognition

Two dimensional (2D) face recognition is a vastly researched topic because of the availability of equipment related to face recognition. Nowadays, 2D cameras are routinely deployed in many different security scenarios. However, 2D face recognition suffers from factors encountered in practice, such as pose and lighting variations, expression variations, age variations, and facial occlusions. Figure 1.2 (K.jain, 2009) shows some different types of problems related to face recognition.

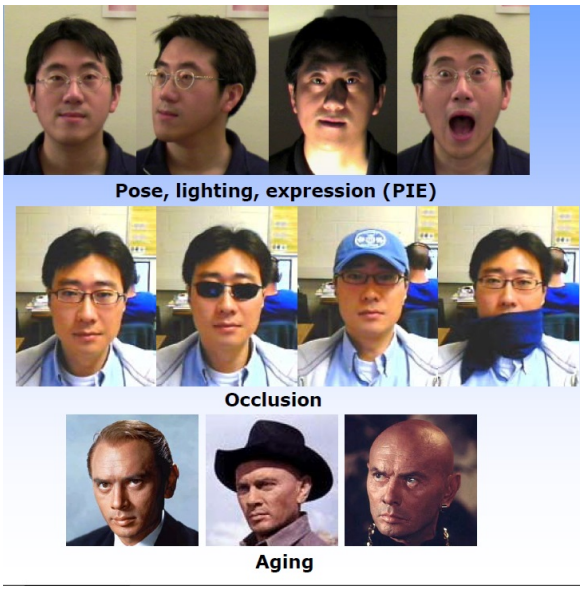


Figure 1.2: Different problems in face recognition

1.2.1 Pose Variation

One major challenge is introduced by pose variation (Bowyer, Chang, & Flynn, 2006; Ding & Tao, 2016). It results in performance degradation of image recognition. The face is a 3D object and the dimensions and distance of its features change drastically if the field of view is changed. Thus two images taken from different views are different

in nature. Thus we have to differentiate not only between between inter-user variation but also intra-user variation in face recognition.

1.2.2 Lighting variation

It is well known that severe lighting variation can also create differences in images drawn from the same person. The face is a 3D object and the lighting source projects itself on a 2D space. Thus different illumination conditions and shades can create different images of the same person. The algorithms that have been developed that are robust to lighting variations exploit prior knowledge of the optics of the lighting sources (Abate, Nappi, Riccio, & Sabatino, 2007; Bowyer et al., 2006).

1.2.3 Occlusion variation

Occasionally, occlusion can result in face change, for example glasses on the face can create variation of features. Some commercial applications reject images if eyes are hidden behind glasses or cannot be detected. We resort to local based methods that ar proposed to overcome the occlusion problem (Martínez, 2002; Tan, Chen, Zhou, & Zhang, 2005).

1.2.4 Expression

Facial expression also causes variation of the face image. For example, a smile can change the facial features. This challenge is met through 3D modeling approaches (Bowyer et al., 2006; Ahonen, Hadid, & Pietikainen, 2006). Although it overcomes to a great extent the variations caused by expressions, it is still an active research area in human computer interaction and communications (Uçar, Demir, & Güzeliş, 2016; Tan, Chen, Zhou, & Liu, 2009)

1.2.5 Age variation

The effects of aging and variations caused by it on the facial features have not been exhaustively studied. Aging related variations can take various forms i) wrinkles and speckles, ii) weight loss and gain and iii) change in face shape primitives (e.g. drooping cheeks, sagged eyes and changes in mouth). Therefore, the aging process affects both shape and texture. All these variations result in degradation of our recognition performance. Moreover, there are only two publicly available face aging databases, viz FG-NET and MORPH. Figures 1.3 and 1.4 show sample images from both databases. These databases exist with all challanges that have been mentioned above. Therefore, it is a challenging task to deal with such databases. As such, facial aging has become an important research problem in the face recognition domain.

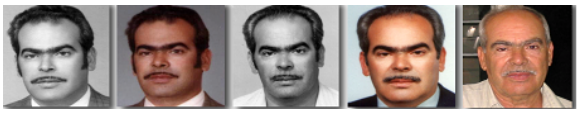


Figure 1.3: Sample of images from FG-NET database

Figure 1.4: Sample of images from MORPH Album 2 database

1.3 Motivation, Objective and Research Questions

1.3.1 Motivation

There are many instances where we have to compare an image with another image that could be a few decades older. This situation arises in many different real life applications, e.g. missing children case, inheritance problems, security and surveillance systems as well as human computer interaction. As yet, limited accuracy is achieved in age invariant face recognition and so there is still room for improvement in accuracy. Another gap in research relates to research focus; all currently proposed techniques only concentrated on increasing recognition accuracy. As such, it would be interesting to formulate methods that would aid in knowledge discovery and understanding relating to the aging process. That represent another objective of this thesis.

In relation to this, the specific objectives of the work conducted in this PhD study are defined as follows:

1.3.2 Objective

The main goal of this research is to develop a novel framework for age invariant face recognition via aging models. The aging models viz global, local, personalized and integrated models will be developed at different levels of data granularity. These aging models would be helpful in understanding the aging process, and as a result, the accuracy of age invariant face recognition is expected to be improved.

Since separate models are developed for shape and texture for a given person, it is possible that two different matching predictions will be produced from each of them,

hence giving rise to a conflict. Thus, another objective of this research is to resolve conflicts between texture and shape-based models. As shown in Chapter 7, these developed models prove to be useful for knowledge discovery and knowledge extraction on the aging process. In this research we have also introduced the novel framework based on NeuCube (SSTD) for gender recognition and age group classification.

1.3.3 Research Questions

The questions that we investigate during this research are.

1. How to develop comprehensive methods for age invariant face recognition that result in better accuracy and effciency than current state-of-the-art methods?

2. How to develop a framework that incorporates both texture and shape features at system level while effectively dealing with conflicts between the two model types?

3. How to develop a craniofacial growth model that maximizes predictive power, as measured by recognition accuracy?

4. Can aging effects be captured at different time points with proposed aging models?

5. How to develop efficient methods for inter-related problems with face recognition, viz age group clasification and gender recognition in order to increase the accuracy?

6. Is knowledge extraction possible with the proposed methods for gender recognition and age group classification?

In order to determine answers to these questions four different types of models built on different levels of data granularity and with different perspective will be investigated.

1.4 Scope

The scope of this research is limited because of time constraint inherent in PhD studies. The face identification problem is carried out by using a closed set testing strategy. This is also true for the gender recognition and age group classification problems. The face age simulation and age estimation is not within the scope of this PhD study.

1.5 Main Contributions

Throughout the completion of the PhD study, seven main contributions have been made:

1. Development of global and personalized models for capturing aging patterns with texture and shape information. Personalized model covers the individual aging patterns while a Global model captures general aging patterns in the database. De-aging factor is introduced that de-ages each individual in the probe set.

2. Development of automatic selection of subspaces with texture and shape information is proposed.

3. Development of a local model for texture and shape information which captures homogeneous sub-population.

4. Development of a multi-model framework named as the Integrated Multi- Model Framework (IMMF) for texture and shape information that integrates different types and levels of aging process and finds the synergy between the three models viz global, local and personalized.

5. Development of Composite model which resolves the conflicts between all four shape and texture models (fourth being integrated model) which would be helpful in more trusted decision making for face recognition.

6. Development of craniofacial growth model which simulates cranfacial growth of face and extracts knowledge from aging models in future time points.

7. Development of NeuCube aging models and use of these models for age group classification and gender recognition from databases and study the effects of aging on genders and age groups.

   In addition a number of publications including journal papers and conference papers have also been produced which are listed below:

1.5.1 Journal Publications

1. Fahad Bashir Alvi & Russel Pears (2016). A composite spatio-temporal modeling approach for age invariant face recognition. Expert Systems with Applications.

2. Fahad Bashir Alvi, Russel Pears & Nikola Kasabov. An evolving spatio-temporal approach for gender and age group classification with Spiking Neural Networks. Evolving Systems.

3. Kasabov, N., Scott, Fahad Bashir Alvi (2016) and others. Evolving spatio-temporal data machines based on the NeuCube neuromorphic framework: design methodology and selected applications. Neural Networks, 78, 1-14.

1.5.2 Conference Publications and Presentations

1. Fahad Bashir Alvi and Russel Pears. Use of Spatio-Temporal Modeling for Age Invariant Face Recognition in 30th International Conference on Image and Vision Computing, New Zealand (IVCNZ 2015).

2. Fahad Bashir Alvi and Russel Pears. An Integrated Modeling Approach to Age Invariant Face Recognition in International Conference on image vision and computing, France (ICIVC 2014).

3. Fahad Bashir Alvi and Russel Pears, Texture Modelling for Age invariant Face Recognition. The 19th International Conference on Image Processing, Computer Vision, & Pattern Recognition Las Vegas, Nevada, USA (IPCV'15).

4. Fahad Bashir Alvi, Russel Pears & Nikola Kasabov. Spatio temporal Data Modelling for Age Classification. Abstract Accepted at 13th International Conference on Neuro Computing and Evolving Intelligence 2015 (NCEI 15).

5. Presentation on Aging Simulation for Face Recognition through Age Simulation in Australian Doctoral Consortium (ACDC'14).

1.6 Thesis Structure

The thesis is structured as follows:

- CHAPTER 1 presents an introduction to the PhD study and a brief description of challenges related to face recognition.

- CHAPTER 2 presents a review of the face recognition and inter-related probelms. It also reviews and explains the features used in this study.

- CHAPTER 3 presents a global and personalized models for capturing aging patterns with texture and shape information. Personalized model covers the individual aging patterns while a Global model captures general aging patterns in the database. De-aging factor is introduced that de-ages each individual in the probe set. It employes KNN and regression for constructing the aging models. It also uses bayesian method for autmomatic selection of subspace. The developed models are then put in place for age invariant face recognition. The publicaly available datasets MORPH and FG-NET were used for evaluating the performance of the proposed models.

- CHAPTER 4 presents a local model for texture and shape information which captures homogeneous sub-population for consturcting the aging models. It employes dynamic time warping and the k means clustering algorithm for constucting the clusters. Then aging models are developed by using nearest neighbour and regression methods. The databases used in Chapter 3 are again being utilised in this chapter to evaluate local model capability for age invariant face recognition.

- CHAPTER 5 presents an Integrated Multi-Model Framework denoted as IMMF proposed by Kasabov (2007b) and Widuputra (2011). The integrated framework utilises the proposed global (Chapter 3), personalized (Chapter 3), and

local (Chapter 4) of aging models for texture and shape information and uses the synergy that exists between the three types of models for age invariant face recognition. The same publicly available MORPH and FG-NET databases are used for evaluation purpose.

- CHAPTER 6 presents the composite model for resolving the conflict between texture and shape for models builts in chapter 3, 4 and 5 at different levels of information. It uses a decision tree method for resolving conflicts between different types of aging models. The same publicly MORPH and FG-NET databases are used for evaulation purposes. It also discusses about different models and explains which model is useful at different situations.

- CHAPTER 7 presents the craniofacial growth model. It also discusses knowledge discovery from aging models at different age points for texture and shape information.

- CHAPTER 8 presents the NeuCube aging models based on spiking neural networks for age group classification and gender recognition . The NeuCube aging models are tested on both pubicaly available datasets.

- CHAPTER 9 presents the discussion and conclusion of the study and suggests future work.

Chapter 2

Background and Literature Review

Aging problem in face image recognition is an important and significant part of the domain of biometric research (A. K. Jain et al., 2016; Cootes et al., 1999; Gong et al., 2015; Li et al., 2011; Ramanathan & Chellappa, 2006; Park et al., 2010). It has important applications in real life in the fields of missing children identification, passport verification, security, animation and business intelligence. For example, in the case of law enforcement, an image of a suspect is available and we need to find out whether the image of the same person exists and can be obtained from a crime database or not. If the suspect does actually appear in the crime database at a previous age then a match should be made and information on the last known address and other associated information of the suspect could possibly be retrieved from the crime database.

Inspite of extensive research in face recognition (Ramanathan & Chellappa, 2006; Gong, Li, Lin, Liu, & Tang, 2013; Gong et al., 2015; Li et al., 2011) much ground has not been covered in the field of age invariant face recognition (Gong et al., 2015, 2013; Li et al., 2011; Park et al., 2010). There are two types of approaches used in age invariant face recognition, generative and non generative. Generative approaches

assume prior knowledge of human age, given an image. On the other hand, non-generative approaches concentrate on finding discriminative features of the face and the changes therein throughout the face aging lifespan. From another perspective we can identify three main directions of research related to face recognition across age: viz, age invariant face recognition, age simulation and age estimation. Most research has concentrated on age estimation and age simulation. Research into age invariant face recognition is still at a nascent stage. There are five major challenges which affect the performance of face recognition systems: pose, illumination, expression, occlusion and aging. It has been observed in past research that for resolving one challenge we need to compromise on others (Geng, Zhou, & Smith-Miles, 2007; Abdullah et al., 2014; Pujol & García, 2012). Thus, in order to find an effective solution to the problem, there is a need to compromise on the minimum number of factors or challenges.

Age estimation means estimating the age of a person from a given image of that individual. On the other hand age simulation deals with the production of images that represent an individual over different time points, taken forwards in time, or backwards in time; the latter corresponding to the situation where a reconstruction of the appearance of the individual at a younger age is obtained. Yet another aspect of face recognition research is age invariant face recognition which deals with locating such features on a face that remain invariant throughout the aging process, or determining a process that annuls the effect of age on those features. It can further be divided in two types, namely generative and non-generative as will be explained below. Figure 2.1 explains these different directions of research.

(Kwon & da Vitoria Lobo, 1999) proposed a classification approach based on age

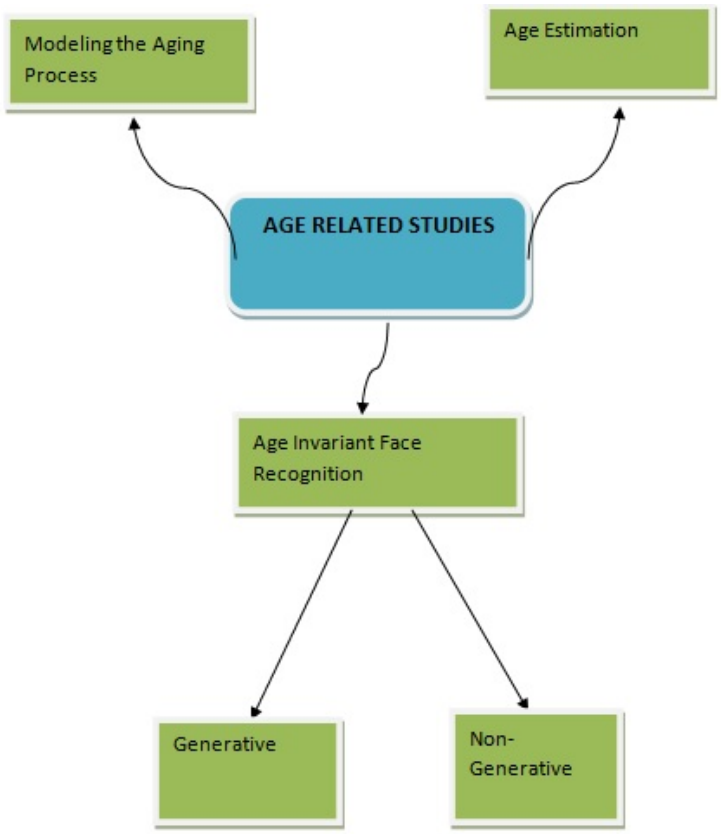


Figure 2.1: Major directions of research

that identifies the age group of given face images. It is based on the anthropometry of the face and the density of its face wrinkles. They considered three age groups in their study. These groups were infants, young adults and senior adults. They also observed that the lower and upper halves of faces grow at different rates during the early formative years. Consequently they used ratios of facial measurements in distinguishing images of infants from those of adults. They identified six ratios for their study. Further, they also proposed the use of snakelets (Kass, Witkin, & Terzopoulos, 1988) in characterizing facial wrinkle density in certain predesignated

facial regions. This was used to distinguish images of young adults from that of senior adults.

(Lanitis, Taylor, & Cootes, 1999) also used Active Appearance Models (AAM). Lanitis et al used a statistical face model for studying the age estimation problem. AAM parameters were extracted from facial images which were marked with 68 points. A Genetic Algorithm was then applied to build and optimize an aging function.

(Geng et al., 2007) worked on the assumption that similar faces age in similar ways across a population. He introduced the method of Aging pattern Subspace (AGES) for modelling aging across time. Geng et al constructed a representative subspace by utilizing a sequence of images of a given individuals face sorted in chronological order. Thereafter, minimum reconstruction error was achieved by projecting the unseen face image into its proper aging subspace, thus making it possible to determine the age of the image from its aging pattern.

(Yan, Zhou, Liu, Hasegawa-Johnson, & Huang, 2008) introduced a method which developed the concept of coordinate patches and Gaussian Mixture Models (GMMs) which were used to estimate facial ages. In their method, the facial image of an individual is encoded as a group of overlapped spatially flexible patches (SFPs). Local features are then extracted by a 2D discrete cosine transform (DCT). The patches are used for integration of local features and coordinate information. The extracted SFPs are then modelled with Gaussian Mixture Models (GMMs). The model produced was used to estimate age of a person by comparing the sum of likelihoods from the set of spatially flexible patches (SFPs) associated with the hypothetic age (i.e. predicted age).

(G. Guo, Fu, Dyer, & Huang, 2008) designed a localized regression classifier to learn the aging function. The concept of manifold learning was used for age estimation. (Fu & Huang, 2008) also used manifold learning as well. A set of age separated images were taken and a low dimensional manifold was developed from it. Linear and quadratic regression functions were applied on the low dimensional feature vectors from the respective manifolds.

(Biswas, Aggarwal, Ramanathan, & Chellappa, 2008) proposed that there is coherency in facial feature drifts across ages. The degree of drift was used as a measure to perform face verification across ages. The assumption was that facial feature drifts observed in images that are age separated follows a definite coherent drift pattern. It was also observed that the same might not be true for age separated face images of two different individuals. Biswas et al extracted fiducial features from the interior region of the face. The points on the outer contour (facial boundary) tend to be very sensitive to head pose variations; therefore a suitable pose correction mechanism had to be put in place.

Ramanathan and Chellappa (Ramanathan & Chellappa, 2006) used eight ratios of distance measures to model age progression in young individuals, ranging in age from 0 to 18 years. These ratios were used to estimate an individual's appearance across time and to perform face recognition across age progression. They experimented with a database of 233 images of 109 individuals, partially collected from the FG-NET aging database and partially from a private gallery of images. Improvments in rank 1 accuracy of 8 percent (without age prediction) to 15 percent over the state of the art approach prevalent at that time (with age transformation) was reported.

(Park et al., 2010) presented an approach to age invariant face recognition by

using a 3D generative model for face aging. In their method, in order to compensate for the age effect, probe face images are first transformed to the same age as the gallery image by using a trained 3D aging model. Then, Face VACS, a commercially available face recognition engine, was used to evaluate the identification results.

Suo et al (Suo, Zhu, Shan, & Chen, 2010) proposed a statistical model for face age simulation and age estimation. A hierarchical ANDOR graph representation is adopted with which faces are decomposed into different parts (e.g., hair, wrinkles, etc.) and organized into graph structures. In each age group faces are represented by means of AND-OR graphs, where the 'AND-nodes' correspond to coarse-to-fine representation of faces and the 'OR-nodes' represent the large diversity of faces by alternative selections. Facial aging was modeled by means of a dynamic Markov process on the AND-OR graph representation. The parameters of the dynamic model are learnt from a large annotated face data set and the stochasticity of face aging is modeled in the dynamics explicitly. The dictionary of different facial components and regions is created across five age groups (20-30 years, 30-40 years, 40-50 years, 50-60 years and 60-70 years).

(Li et al., 2011) proposed a discriminative model to address face matching in the presence of age variation. Each face was represented by designing a densely sampled local features description scheme, in which scale invariant feature transform (SIFT) and multi-scale local binary patterns (MLBP) were used as the local descriptors. The densely sampled features were provided sufficient discriminatory information, including distribution of the edge direction in the image. The multi feature discriminant analysis method was developed to avoid over fitting. Multi feature discriminant analysis combined two different random sampling methods, operating in feature and sample

space. Finally, a fusion rule was used to classify by constructing multiple LDA-based classifiers. Experiments were conducted with FG-NET and MORPH databases and 47.5% and 83.5% accuracy were obtained respectively.

(Syambas & Purwanto, 2012) used a combination of Active Appearance Models (AAMs), Support Vector Machines (SVMs) and Monte-Carlo simulation to build a high accuracy aging model. Two types of experiments were reported. In the first experiment, they showed empirically that as the probe face progresses in age, the face recognition rate decreases. In order to overcome this problem, in a second experiment, they artificially aged probe faces so as to increase the likelihood of a match with a gallery image that would then be closer in time terms to the aged version of the probe image. The artificial aging was accomplished by the use of active appearance models while the matching process was based on the use of PCA. A rank 1 accuracy of 32% was obtained by Sethuram et al.

(Gong et al., 2013) proposed a Hidden Factor Analysis (HFA) method in which person-specific stable features are extracted by analyzing the variation caused by the aging process. A probabilistic model was built with two latent factors: an identity factor that is age invariant, and an age factor affected by aging process were used to separate the variation caused by aging process. Thereafter, the two latent factors were combined in order to model the appearance. An expectation maximization procedure was also developed to jointly estimate the latent factors and the model parameteres. The proposed methods were tested with FG-NET and MORPH database

Gong et al (Gong et al., 2015) proposed a new method for face representation and matching for the age invariant face recognition problem. A new maximum entropy feature descriptor (MEFD) method is developed that encodes the microstructure of

facial images into a set of discrete codes. The code entropy is maximized in order to extract discriminative and expressive information from densely sampling encoded face images. An identity factor analysis method was developed to estimate the probability that two given faces have the same underlying identity. The method was tested on the FGNET anad MORPH datasets and accuracies of 76.2% and 92.26% respectively were obtained.

As mentioned above, Age invariant Face recognition methods can also be classified as generative or non-generative. Recognition methods under the Generative approach category typically introduce appearance transformations on one of the test images to reduce the facial appearance difference due to age separation. These approaches typically involve a computational model for facial aging which is subsequently employed on a face recognition task. The face recognition tasks are described in (Lanitis et al., 1999; Ramanathan & Chellappa, 2006; Park et al., 2010).

On the other hand, non-generative approaches derive an age invariant signature from faces which is used to perform face recognition across age progression. The approaches proposed by (Ling, Soatto, Ramanathan, & Jacobs, 2007; Li et al., 2011; Gong et al., 2013, 2015) fall under the latter category.

Tables 2.1 and 2.2 summarize accuracy achieved by various prominent approaches on the FG-NET and MORPH datasets respectively.

Table 2.1: Showing Prominent Results with FG-NET database

Models	**Accuracy**
(Geng et al., 2007)	38.1%
(Park et al., 2010)	37.4%
(Li et al., 2011)	47.50%
(Gong et al., 2013)	60%
(Gong et al., 2015)	76.2%

Table 2.2: Showing Prominent Results with MORPH database

Models	**Accuracy**
(Park et al., 2010)	79.80%
(Li et al., 2011)	83.90%
(Gong et al., 2013)	91.14%
(Gong et al., 2015)	94.59%
(Li, Gong, Li, & Tao, 2016)	94.87%

2.1 Discussion

On the basis of the above literature review, the design of an appropriate feature representation and an effective matching framework for age invariant face recognition remains an open problem. It is concluded that aging of a human face results in both change of texture as well as in shape information. This research is an attempt to capture the change, study its pattern and then utilize it for modelling the aging phenomenon. In order to achieve highest possible accuracy there is a need to study both types of features, ie the shape and the texture information of the human face.

The Anthropometric model has the potential of giving detailed information regarding changes in the geometry of the face. Consequently, it was decided to extract shape representation of face with the help of the Anthropometric model. Ratios of the geometical distances between various points on the face were utilized to create shape-based facial feature. As the face ages, the skin texture changes color and develops wrinkes. These wrinkes which are in the form of lines are also very helpful in characterizing the aging process. It was also noted in the literature that in earlier ages, changes in shape are more significant when compared to texture changes, while in the older ages, the opposite is true (Suo et al., 2010; Ramanathan & Chellappa, 2006; Rowland & Perrett, 1995; Bruce & Young, 1986; T. Wu, Turaga, & Chellappa, 2012; Mondloch, Le Grand, & Maurer, 2002): the texture and wrinkles on the face provide more information about aging in older age groups. It was decided that for study of texture the information and data of edges occurring on the face should be gathered and treated as local features of the face. As a person ages, the number of edges on the face will increase, thus providing data for the aging process. It was decided to use the canny edge detection technique (Tian, Kanade, & Cohn, 2000; Gheissari, Sebastian, & Hartley, 2006; Jana, Datta, & Saha, 2013; Ding, Choi, Tao, & Davis, 2016)

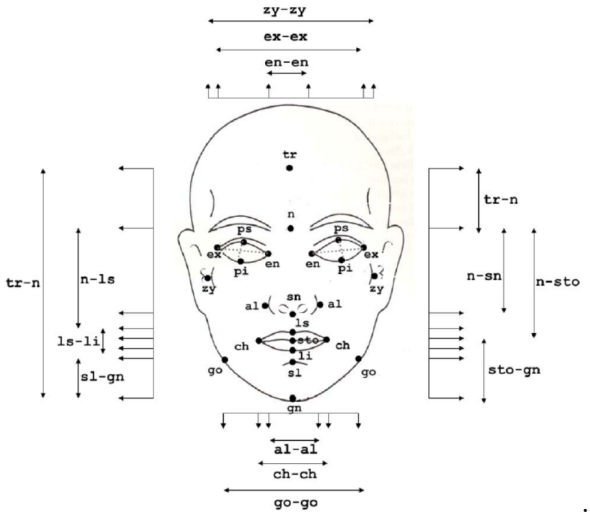
2.2 Shape Features

2.2.1 Anthropometric Model

Face anthropometry, the science of measuring sizes and proportions on human faces, has the potential to play a crucial role in developing facial aging models. Such studies provide a quantitative description of the craniofacial growth at different ages and

hence provide a plethora of options for learning based approaches to be adopted to characterize facial aging. Face anthropometric studies provide dense measurements taken between key landmarks on human faces across different ages and have played a critical role in surgical procedures employed on the faces of growing children (Farkas & Munro, 1987). Farkas provides a comprehensive overview of face anthropometry and its many significant applications. He defines face anthropometry in terms of measurements taken from 57 carefully selected landmarks on human faces spread across 6 regions in the craniofacial area (head, face, orbits, nose, lips and mouth, ear). The facial measurements are of three kinds: (i) projective measurements (shortest distance between two landmarks); (ii) tangential measurements (distance between two landmarks measured along the skin surface) and (iii) angular measurements. Figure 2 illustrates the kind of data that is collected in face anthropometric studies and further illustrates the different fiducial features across which such data is collected (Farkas, 1994). In the absence of age-based Anthropometric measurements they collected facial growth data by extracting facial features on the passport database.

Such growth data was collected on five different age groups: 21-30 years, 31-40 years, 41-50 years, 51-60 years and 61- 70 years. The facial growth data collected in this manner was found to be effective in characterizing facial growth based on age, gender, ethnicity etc. and further during cases where individuals gain or lose weight.



- Facial index ($\frac{n-gn}{zy-zy}$)
- Mandibular index ($\frac{sto-gn}{go-go}$)
- Intercantal index ($\frac{en-en}{ex-ex}$)
- Orbital width index ($\frac{ex-en}{en-en}$)
- Eye fissure index ($\frac{ps-pi}{ex-en}$)
- Vermilion height index ($\frac{ls-sto}{sto-li}$)
- Mouth Face width index ($\frac{ch-ch}{zy-zy}$)

Figure 2.2: Anthropometric Model and Seven Features.

2.3 Texture Features

In the case of texture, as a part of preprocessing, we first register(Goshtasby, 2012; Štruc & Pavešić, 2010) images on the basis of eye coordinates so that all images transform to the same size and that each point on any given face refers to the same point on all of the images. We then divide the facial image horizontally into five slices. The slices cover the forehead, eyes, nose, mouth and the area below the mouth. Edges are marked on the face after converting the image into grayscale. A histogram is prepared for these edges. The frequency of the edges in each slice is taken as a feature, thus resulting in five features across the entire face as shown in Figure 2.3.

Figure 2.3: Five slices of the face.

2.3.1 Edge Detection

We used the Canny edge detector to find edges. The Canny operator was designed to be an optimal edge detector (according to particular criteria — there are other detectors around that also claim to be optimal with respect to slightly different criteria). It takes as input a gray scale image, and produces as output an image showing the positions of tracked intensity discontinuities.

The Canny edge detector operator works as a multi-stage process. First of all the image is smoothed by Gaussian convolution. Then a simple 2-D first derivative operator is applied to the smoothed image to highlight regions of the image with high values for the first spatial derivative. Edges give rise to ridges in the gradient magnitude image. The algorithm then tracks along the top of these ridges and sets to zero all pixels that are not actually on the ridge top so as to give a thin line in the output. This process is known as non-maximal suppression. The tracking process exhibits hysteresis and is controlled by two thresholds: $t1$ and $t2$, with $t1 > t2$. Tracking can only begin at a point on a ridge higher than t1. Tracking then continues in both directions out from that point until the height of the ridge falls below $t2$.

.

Figure 2.4: Canny edge detection.

This hysteresis helps to ensure that noisy edges are not broken up into multiple edge fragments (Canny, 1986). Figure 2.4 shows canny edge detection for a selected image.

Chapter 3

Global and Personalized Modelling

3.1 Introduction

This chapter introduces two novel aging models, viz Global and Personalised models, each one based both on texture and shape information. A Global model is a distinct, fixed and reusable model that is very useful for describing the general underlying behaviour of a stochastic system. This model is a realisation of inductive reasoning that builds a single model by learning from the entire data set or problem space. The developed model is then applied to new data that arrives in the future (Baruch & Stoyanov, 1995; Spillantini et al., 1997; Marin, Garcia-Lagos, Joya, & Sandoval, 2002). Global modelling is the most commonly-used approach for inductive reasoning. So, this model investigates the general trend in aging. The Personalised model investigates a given individual's aging trend. Every person ages in a different way that is unique to that person. Therefore Personalised modeling is required in order to capture the personalised aging process.

3.2 Computational Models for Aging

In this research we treat the age recognition problem as a time series problem with each texture feature and anthropometric feature giving rise to a separate time series variable. In the standard time series problem a trajectory at several different time points from the past is available and the problem is to project into the future. The time series problem that we tackle differs from the standard time series problem in one fundamental aspect: in the face recognition scenario only one time point is available, which is the image (feature vector) at the time point at which the image was captured. No other training data is available. This would make the problem virtually unsolvable if not for the fact that we can explore the data spatially to obtain more training data. Thus at any given time point we have f data points obtained from the f different anthropometric features that we extract. However, since the time series trajectory of each feature is largely independent of the others, a simple approach of combining them into a single series would be highly ineffective. As such, our solution is to first build trajectories for each of the features separately and then combine predicted feature values at each time point into a single vector which is used to determine whether a match exists between probe and gallery images. The other relatively minor difference is the direction in which the trajectory is built; in our case it is backwards in time instead of forwards.

In this chapter we investigate two different types of models built at different levels of data granularity. At the global level a model is built on training data that encompasses the entire set of available individuals. At the individual level a personalized model was built for each individual. In the evaluation stage we use a similarity matrix to compute Rank 1 accuracy.

3.2.1 Global Model

Model Construction

A global model is induced over the entire problem space that contains the entire population of available individuals. Such models serve to capture useful general trends across the population over a spectrum of age bands that we use. After grouping images into their respective age bands, we determine discriminative features and compute the values of 7 unique indexes (i.e. ratio of Euclidean distances for a given pair of Anthropometric features) for each image. The indexes were calculated after warping the image to a mean image using the Procrustes algorithm (Cootes et al., 1999). In the case of texture, as a part of preprocessing, we first register the images, so that all images transform to the same size and each point on any given image refers to the same point on all of the images. Then we divide the facial image into five horizontal slices. The slices cover the forehead, eyes, nose, mouth and the area below the mouth. Edges(Canny) are marked on the face after converting the image into grayscale. A histogram is prepared for these edges. The frequency of edges in each slice is taken as a feature, thus resulting in five features for the whole face. The centroid of each age band is determined and an n^{th} order polynomial (we experiment with values of n in the range [1..3]) function is developed that spans all the age bands using a least square based non-linear regression method. Equation (3.1) represents the Global model for the Facial index feature, where c_i represents the value of each coefficient and x_i represents its age band index value. Figure 3.1 shows the trajectory of the Global Model across the agebands for the Facial index.

$$y_i^{(global)} = c_1 x_i^n + c_2 {x_i}^{n-1} + ... + c_n x_i + c_{n+1} \tag{3.1}$$

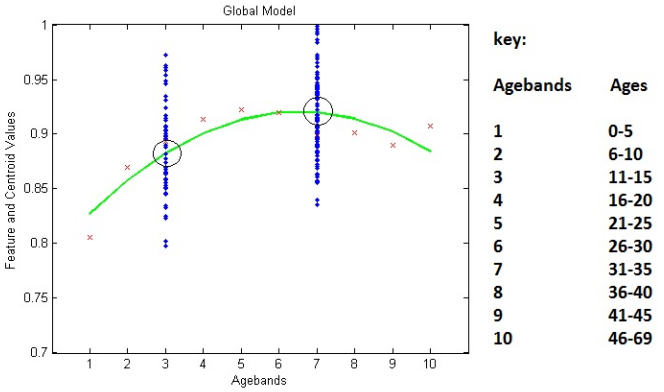


Figure 3.1: Global Model for Feature 1 (Facial index). The two vertical blue lines show the spread of feature values in age band 3 and age band 8.

3.2.2 Personalized Model

Model Construction

The transductive or personalized approach, in contrast to the inductive approach, models each point in the problem space. It was defined by Vapnik in (Vapnik & Vapnik, 1998) and used by Kasabov in (N. Kasabov, 2007b) and Pears in (Pears, Widiputra, & Kasabov, 2013). The intuition behind personalized modeling is that the aging process differs from person to person and hence modeling at the level of individuals can be expected to yield more accuracy. The k-NN (k-nearest neighbor) is one of the well-known transductive techniques and is the most widely used form of personalized modeling.

The model is constructed as follows. We take an image, determine its age band (in closed set evaluation mode the age for the image is known; in real-world applications, a human expert can be used to estimate the age) and obtain its k nearest neighbors (the

optimal value of k=5 is set by experimentation). When we get an Anthropometric sample of an image we de-age that sample as follows:

$$DF_i = M_i - F_i \tag{3.2}$$

We introduce a deviation factor *(DF)* as defined in equation (3.2) above. The deviation factor captures the degree to which an individual's Anthropometric feature value differs from the population as a whole. In this research we assume that the DF_i is independent of age which amounts to assuming that the rate of change in a feature value is constant over time. In equation (3.2), M_i represents the mean of feature i at the age band j estimated by the human expert and F_i is the feature value of the image in that same age band.

$$P_{ij} = G_{ij} + DF_i \tag{3.3}$$

The de-aging process is done in equation (3.3) by adding the deviation factor DF_i of the image in the given age band to the centroid value G_{ij} returned by the global model at age band j which in turn is obtained by applying equation (3.1) on feature i. The resultant pivot value P_{ij} is used to find the k nearest neighbors and determine the centroid a_i of these neighbors. This process is repeated for each age band to complete the construction of the personalized model for the image. We use these centroids a_i to fit a non-linear function across the age bands; once again we fit an n^{th} degree polynomial curve using the method of least squares. Equation (3.4) represents

the Personalized model where y_i is the feature value for feature i; x_i is the age band index and n is the degree of the polynomial.

$$y_i^{(personalized)} = c_1 {x_i}^n + c_2 {x_i}^{n-1} + ... + c_n x_i + c_{n+1} \tag{3.4}$$

Equation (3.4) embodies the aging trajectory for an individual and provides insights into how one individual's aging trajectory differs from another in the given population. Although Global model is used for calculating the deviation factor, it is a personalized process because equation 3.2 and equation 3.3 are used for building separate trajectories for feature i in the backward direction. In our experimentation in sections 3.5 and 3.6 we demonstrate the effectiveness of the personalized model against the global model.

3.3 Use of Subspace for Target Image Identification

In order to improve rank 1 accuracy, we aggregated adjacent age bands into subspaces. In effect, we combine two adjacent age bands into one subspace, thus building I subspaces from N age bands. This has two advantages. Firstly, we expect rank 1 accuracy to improve as there is less uncertainty in locating subspaces as opposed to age bands. This is due to the fact that subspaces are larger in size than age bands. Secondly, image recognition time is reduced as the search is restricted to a smaller number of time segments.

Naive Bayes is used to identify a subspace given a probe feature X_j. Naive Bayes is an effective and computationally efficient technique for constructing a classifier.

The classifier assumes that features are independent of each other which implies that the conditional probability of a class (in our case, a subspace) on a given feature is independent of the corresponding conditional probabilities across the other features. Although this assumption may not hold strictly in practice several empirical studies in the machine learning literature have shown that this type of classifier is robust in such situations as well (Witten & Frank, 2005). We thus select the Naive Bayes classifier on account of its robustness and computational efficiency.

Each subspace is constructed as shown in Tables 3.2 and 3.4. Two bins, *high* and *low*, spanning the higher and lower end of the median boundary range, were used for each subspace after normalizing each feature value to the [0..1] range. When a probe image is taken, models at the global and personalised data granularity levels are constructed. The feature values are de-aged to the age band of the probe image. Then a predicted value for all of the features are produced by the model. These values are compared with the bin values and a bin is chosen on the basis of the closest match with the predicted value.

A posterior probability for each subspace is computed by using equation (3.5) where $Pr(A_i)$ is a prior probability for subspace i. The conditional probability $Pr(X_j|A_i)$ is then found for each feature j by applying our aging models. The probability *Pr(X)* is a scaling factor that is computed across the entire feature set.

$$Pr(A_i|X_j) = Pr(A_i)Pr(X_j|A_i)/Pr(X) \tag{3.5}$$

where

$$Pr(X_j|A_i) = Pr(X_1|A_i)Pr(X_2|A_i).....Pr(X_n|A_i)$$

$$Pr(X_j|A_i) = \prod_{j=1}^{n} Pr(X_j|A_i)$$

$$Pr(A_i|X_j) = Pr(A_i) \prod_{j=1}^{n} Pr(X_j|A_i)$$

$$Pr(X) = \sum_{i=1}^{k} Pr(A_i) \prod_{j=1}^{n} Pr(X_j|A_i)$$

This process is repeated for every subspace. The subspace that yields the highest posterior probability $Pr(A_i|X_j)$ is chosen and is assumed to contain the desired probe image. We used a leave-one-person-out strategy to train the classifier.

For recognition, given a probe image, its age band is obtained from the human expert. We use this age band estimate to compute predicted feature values for the probe image by using the aging models. We then determine the target subspace of the probe image. Finally, a similarity matrix is computed for measuring the performance. The similarity matrix is a numerical measure of how similar two faces are.

A similarity matrix of scores is computed based on the Euclidean distances between all pairs of biometric samples from the gallery and the probe p_j image. The similarity matrix is built from all images in the search space and rank 1 accuracy is then computed. The identification rate at rank 1, $P_I(1)$, is also called the correct identification rate and the top match rate or score (A. K. Jain & Li, 2011; A. Jain et al., 2007). The identification rate for rank n, $P_I(n)$, is shown in equation (3.7).

For rank n, let

$$C(n) = |\{p_j : rank(p_j) \leq n\}| \tag{3.6}$$

be the cumulative count of the number of probes at rank n or less. The identification rate $P_I(n)$ at rank n is then given by:

$$P_I(n) = \frac{|C(n)|}{|P_G|} \tag{3.7}$$

where $|P_G|$ is the number of images in the probe set. Both $C(n)$ and $P_I(n)$ are non-monotonic decreasing functions of n.

In closed set system, CMC curve is used for evaluating the performance of face recognition system where CMC curve depicts identification rates over candidate list sizes. The candidate list is selected by including the top X ranked results(Dunstone & Yager, 2008). The performance metric of interest is where the genuine match is ranked. Instead of asking Is the top match correct?, the CMC plot answers the question, Is the correct match in the top-n matches?(A. K. Jain & Li, 2011; A. Jain et al., 2007).

3.4 Empirical Study

Experiments were performed on the publicly available FG-NET and MORPH Album 2 (the largest publicly available face aging dataset) (Ricanek Jr & Tesafaye, 2006), both of which are used for benchmarking new methods. The lack of a large publicly available face aging database until recently limited research on age invariant face recognition.

There are two desired attributes of a face aging database: (i) large number of subjects, and (ii) large number of face images per subject captured at many different age points (Li et al., 2011). In addition, it is desired that these images should not have large variations in pose, expression, and illumination.

Each of the two datasets that we experimented with has their own challenges. The

MORPH dataset has a large number of subjects with images taken across a narrow age timeline while FG-NET database has a smaller number of subjects and images but the average age gap between images from the same person is much larger than with the MORPH dataset.

The MORPH dataset contains about 55,000 face images from 13,000 different people. The FG-NET database on the other hand contains 1002 color and gray face images of 82 persons across a range of different ethnicities. There is a large variation in lighting, expression and pose across the different images. The image size is 300×400 in pixel units, on the average. The ages vary from 0 to 69 years. There are on the average 12 images per person across different ages. The database was divided into ten different age bands as shown in Tables 3.1 and 3.3.

Table 3.1: FG-NET Age Bands

Ages	0-5	6-10	11-15	16-20	21-25	26-30	31-35	36-40	41-45	46-69
Images	233	178	164	155	81	62	38	31	26	34
Subjects	75	70	71	68	46	38	30	24	19	10

Table 3.2: FG-NET Subspaces

Ages	0-10	11-20	21-30	31-40	41-69
Images	411	319	143	69	60
Subjects	77	81	50	35	22

Table 3.3: Morph Age Bands

Ages	16-20	21-25	26-30	31-35	36-40	41-45	46-50	51-55	56-60	61-65
Images	2189	2113	1642	1760	1825	1657	1018	549	183	64
Subjects	973	1037	821	903	971	850	536	287	92	30

Table 3.4: Morph Subspaces

Ages	16-25	26-35	36-45	46-55	56-65
Images	2283	1759	1680	675	103
Subjects	2283	1759	1680	675	103

3.5 Experiment with FG-NET Database

For evaluation we chose all 1002 images taken from the entire set of 82 individuals. A leave-one person-out (LOPO) strategy was used for evaluation (Gong et al., 2013, 2015; Li et al., 2011).

Our experimentation is focused on comparing the performance of the Global and Personalized models.

3.5.1 Success rate of subspace selection

Table 3.5 shows that the Bayesian classifier was highly effective at finding the correct subspace for both shape and texture models, with both types of models returning

accuarcy rates around the mid 90% mark as measured according to the LOPO strategy mentioned earlier.

Table 3.5: Classification Accuracy of subspace identification

Models	**Database,subjects and Images in probe**	**Accuracy**
Shape	FG-NET(82,1002)	96.4%
Texture	FG-NET(82,1002)	94.6%

3.5.2 Comparison of models across data granularity level for shape features

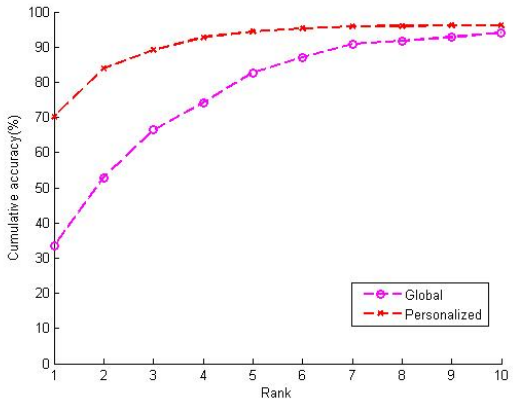


Figure 3.2: Cumulative Matching Characteristic Curves (CMC) for FG-NET Database

The CMC curves in Figure 3.2 shows the performance of the Global and Personalized models on shape features across rank orders from 1 to 10, with the Personalized model

emerging winner as the clear winner across the range.

3.5.3 Comparison of models across data granularity level for texture features

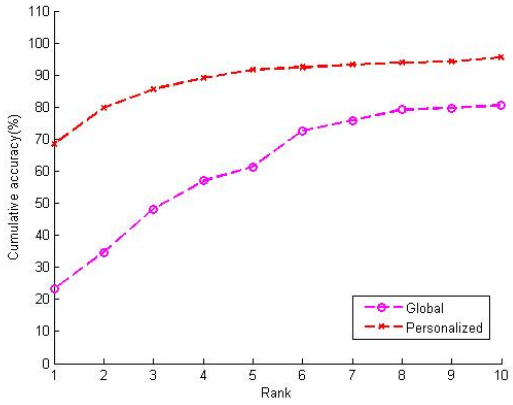


Figure 3.3: Cumulative Matching Characteristic Curves for FG-NET Database

The CMC curves in Figure 3.3 for texture features once again show that the Personalized model is the clear winner across across rank orders from 1 to 10.

We note that the models developed can not only be used for gallery image extraction but could be used to simulate the aging process as well. The age trajectory functions that we build can be used to predict their features values by performing a forward traversal (rather than a backward one). This a further advantage of our aging models over the one proposed by (Gong et al., 2015).

Table 3.6: Comparative results of Rank-1 score with FG-NET database

Models	**shape**	**Texture**
Global	33.7%	23.3%
Personalized	70.4%	68.6%

Table 3.7: Comparative results of Rank-1 score with FG-NET database

Models	**Rank-1**
Global (Texture)	23.3%
Global (Shape)	33.7%
Personalized(Texture)	68.6%
Personalized(Shape)	70.4%
(Geng et al., 2007)	38.1%
(Park et al., 2010)	37.4%
(Li et al., 2011)	47.5%
(Gong et al., 2013)	69%
(Gong et al., 2015)	76.2%

3.6 Morph Experimentation

In order to check generality of our models on different types of databases, we used MORPH Album 2 (the largest publicly available face aging dataset) database. We annotated the fiducial landmarks with Stasm as shown in Figure 3.4, a method based on Active shape Models (Milborrow & Nicolls, 2014). We used all 13000 subjects for conducting the experiment. We used the same methodology as used in state of art algorithms (Gong et al., 2013, 2015; Li et al., 2011) in which the MORPH album 2 data set was split into three sets, viz training set, a probe set and a gallery set.

Figure 3.4: Fiducial landmark detection results. Blue points represent the 68 landmark points

The persons in the training set were entirely different from those in the probe and gallery sets. For the training dataset, we selected a subset of 13,000 face images from 6500 subjects, with two images per subject. These two images were selected such that they had the largest age gap between them. The gallery set was composed of 6500 face images corresponding to the youngest age of these 6500 subjects. The probe set was composed of 6500 face images corresponding to the oldest age of these 6500 subjects.

The CMC curves in Figure 3.5 and Figure 3.6 confirm the trends shown in Table 3.8 across rank order, thus confirming that the Personalized model is once again the overall winner on the MORPH dataset as well.

3.6.1 Success rate of subspace selection

Table 3.7 shows that the Bayesian classifier was highly effective at finding the correct subspace for both shape and texture models, with both types of models returning accuracy rates above 95% as measured according to the LOPO strategy mentioned earlier.

Table 3.8: Comparative results of Classification of different Models

Models	**Database,subjects and Images in probe**	**Accuracy**
Shape	Morph(6500,6500)	98.8%
Texture	Morph(6500,6500)	99.0%

3.6.2 Comparison of models across data granularity level for shape features

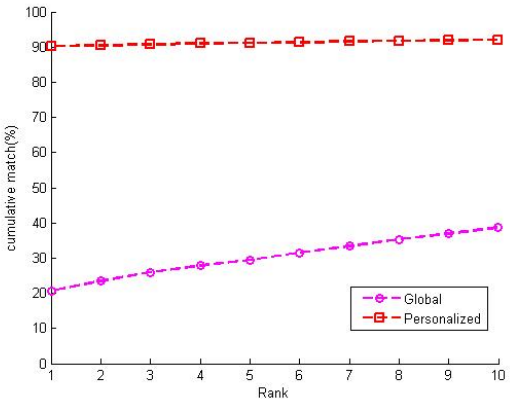


Figure 3.5: Cumulative Matching Characteristic Curves for Morph Database

The CMC curves in Figure 3.5 confirm that the trends displayed in shape models hold across rank orders from 1 to 10, with the Personalized model emerging the winner across the range.

3.6.3 Comparison of models across data granularity level for texture features

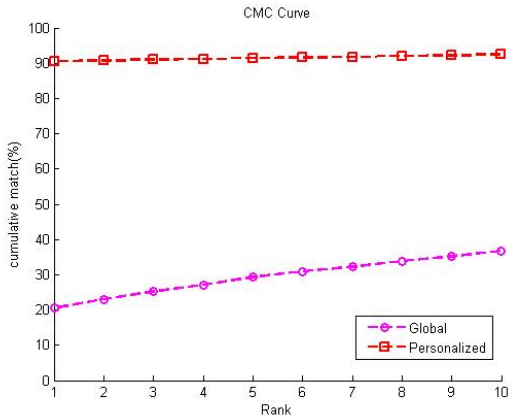


Figure 3.6: Cumulative Matching Characteristic Curves for Morph Database

The CMC curves in Figure 3.6 again confirm that the trends displayed in texture models hold across rank orders from 1 to 10, with the Personalized model emerging the winner across the range.

Table 3.9: Comparative results of Rank-1 score with MORPH database

Models	**Shape**	**Texture**
Global	21%	20%
Personalized	91%	90.5%

Tables 3.6 and 3.9 show that the Personalised Model outperforms the Global Model.

Table 3.10: Comparative results of Rank-1 score with MORPH database

Models	**Rank-1**
Global(Texture)	20%
Global(Shape)	21%
Personalized(Texture)	90.5%
Personalized(Shape)	91%
(Klare & Jain, 2011)	79.08%
(Park et al., 2010)	79.80%
(Li et al., 2011)	83.90%
(Gong et al., 2013)	91.14%
(Gong et al., 2015)	94.59%
(Li et al., 2016)	94.87%

In this set of experiments we investigated the effects of shape and texture models at global and personalized level. In tables 3.7 and 3.10 we benchmarked our models against the current state-of-the-art models. The results clearly indicate that our Personalized model outperformed the models of (Li et al., 2011; Park et al., 2010) but near to models of (Gong et al., 2015). The results in the subsequent chapters 5 and 6 would clearly indicate that our integrated model for shape and texture features and composite model (combination of shape and texture features) have higher Rank 1 accuracy than with the Gong et al approach.

3.7 Discussion

In this chapter, the models of shape and texture at different levels of granularity have been presented. It was observed that dividing the whole database into subspaces and using the Naive Bayesian approach to narrow down the search space proved to be a success. This limits the search space, which reduces the computation cost and

improves the recognition accuracy. The Global model has some limitations because it captures the general behavior of population, thus resulting in low accuracy. Even though the global approach operating on its own was outperformed by personalized model it was indispensable as the personalized approach required a globalized aging function to de-age features and step backward through the age bands. The Personalised model provides very high accuracy. However it has a limitation in that it would be expensive to build a Personalized model for every member specially when the database size is large. Moreover, k nearest neighbor approaches would become more expensive as database size increases. The solution to this problem is to find homogeneous groups in database on the basis of aging pattern. This search lead to development of a local model. This model, built from homogeneous population, would be more helpful for finding probe image thus reducing the computation cost. The next chapter discusses the Local model and building homogeneous population for that model. It is worthwhile to point out here that reduction in size of search space is a different phenomenon from building a homogenous population, as will be evident from the subsequent discussion in the next Chapter.

Chapter 4

Local Modelling

4.1 Introduction

Local models (N. Kasabov, 2007b, 2007a, 2001; Yamada, Yamashita, Ishii, & Iwata, 2006; Lucks, Oki, & RamirezAngulo, 1999; Widiputra, Pears, & Kasabov, 2011; Widiputra, 2011; Hwang, 2009) break down a problem space into smaller subspaces before building separete models on each of the subspaces defined. These models are created by grouping together data that shows similar behaviour within the group while being distinct from the behavior of data outside the group. For example, the variables in a given subpopulation change in consonance with each other but in a different manner from those in other subpopulations. Data for such items of similar behavior define their own clusters. These clusters are the basis of the local models that are developed based on regression. Such models display higher accuracy over the local problem space covered as compared to that of a global model. In this type of modeling, results are evaluated on only a subset of the problem space. Local modeling provides us additional flexibility as results can be predicted either on the basis of single model or in combination with other models (N. Kasabov, 2007a). There

is another distinct advantage in that local models enable us to capture recent data trends in the data and relate them to similar behaviour from the past data. A global model considers only the past activity, thus resulting in reducing the significance of recent trends in the data (Widiputra et al., 2011a).

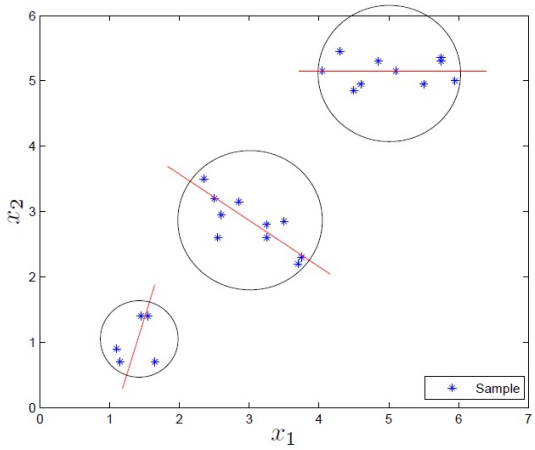


Figure 4.1: Illustration of local modelling to create localised linear regression models from clusters of sample data set in a 2-D space (Hwang,2009, Widiputra et al., 2011)

4.2 Local Model

While a personalized model captures aging effects that are tailored to an individual and a global model captures trends across a population, scope exists for an intermediate level modeling approach. Given a probe image the local modeling approach constructs a model from a homogeneous subset of the population that is most closely aligned with the image rather than from the entire population that in general is heterogeneous. Essentially, the local approach constructs subpopulations through the use of a clustering scheme. For each probe image, the cluster that is closest to it

is identified and thereafter the same model construction process as discussed in the previous chapter for the personalized model is applied. Thus in terms of performance, the local approach is expected to be faster than personalized modeling as the construction and fitting of the aging function is restricted to a subset of instances rather than to the population as a whole.

In a real world situation, the population may be heterogeneous with several different sub-populations made up of different ethnicities, different lifestyles, etc, all of which have a bearing on the aging process. However, in this research we do not assume that the image gallery is annotated with such explicit feature information. Thus in order to segment the population we need to make use of a clustering algorithm that could be applied on the time series trajectories obtained from the fiducial features taken over time. The intuition is that individuals who age similarly will produce similar trajectories and hence will be clustered together in the same segment. However, it can also happen that any two given individuals age similarly but have pair-wise differences in the time series variable (fiducial feature) over the age bands that we track. In time series analysis this corresponds to two time series that are highly correlated but are out of phase with each other. A standard clustering algorithm such as K-means will fail to cluster such correlated but out of phase sequences. This problem can be avoided by applying Dynamic Time Warping (DTW)(Sakoe & Chiba, 1978; Müller, 2007; Sakoe, Chiba, Waibel, & Lee, 1990) before applying K-means. As mentioned earlier the construction follows the same procedure as for the personalized model but with segmentation performed as a pre-processing step and the use of cluster means in place of global means at each age band.

4.3 Dynamic Time Warping

Dynamic Time Warping (DTW) is a well-known technique to find an optimal alignment between two given (time-dependent) sequences under certain restrictions.

The objective of DTW is to compare two (time-dependent) sequences $X = (x_1, x_2, ..., x_N)$ of length $N \in N$ and $Y = (y_1, y_2, ..., y_M)$ of length $M \in N$. These sequences in our case are aging image attributes sampled at equidistant points in time. Let us denote our feature space with F. Then $x_n, y_m \in F$ for $n \in [1 : N]$ and $m \in [1 : M]$. To compare two different features $x, y \in F$, a local cost measure needs to be computed. It is referred to as a local distance measure.

Typically, $c(x, y)$ is small (low cost) if x and y are similar to each other, otherwise $c(x, y)$ is large (high cost).

The algorithm starts by building the distance matrix $C \in \mathbb{R}^{N \times M}$ representing all pairwise distances between X and Y. This distance matrix is called the **local cost matrix** for the alignment of two sequences X and Y:

$$C_l \in \mathbb{R}^{N \times M} \ : \ c_{i,j} = \|x_i - y_j\| \, , \ i \in [1 : N], \ j \in [1 : M] \tag{4.1}$$

Once the local cost matrix is built, the algorithm finds the **alignment path** which runs through the low-cost areas - "valleys" on the cost matrix, Figure 4.2. This alignment path (or **warping path**, or **warping function**) defines the correspondence of an element $x_i \in X$ to $y_j \in Y$ following the boundary condition which assigns first and last elements of X and Y to each other.

4.2. DTW can be formally defined as an (N, M) warping path is a sequence $p = (p_1, p_2, ..., p_K)$ with $p_l = (p_i, p_j) \in [1 : N] \times [1 : M]$ for $l \in [1 : K]$ which must satisfy to the following criteria:

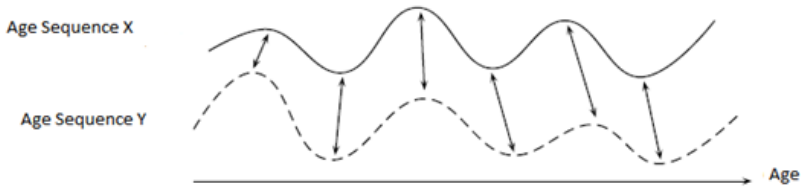


Figure 4.2: Time alignment of two independent time sequences. Alined points are indicated by the arrows

1. **Boundary condition**: $p_1 = (1, 1)$ and $p_K = (N, M)$. The starting and ending points of the warping path must be the first and the last points of aligned sequences.

2. **Monotonicity condition**: $n_1 \leq n_2 \leq ... \leq n_K$ and $m_1 \leq m_2 \leq ... \leq m_K$. This condition preserves the time-ordering of points.

3. **Step size condition**: the basic step size condition formulated as $p_{l+1} - p_l \in \{(1, 1), (1, 0), (0, 1)\}$.

The total cost of the warping is computed as it is used in cluster construction. The **cost function** associated with a warping path is computed with respect to the local cost matrix (which represents all pairwise distances)

$$c_p(X, Y) = \sum_{l=1}^{L} c(x_{n_l}, y_{m_l}) \tag{4.2}$$

The warping path which has a minimal cost associated with alignment is called the **optimal warping path** P^*. Once computed, it used for time-series alignment.

In order to determine an optimal path p, there is a need to test every possible warping path between X and Y . Such a procedure, however, would lead to a computational complexity that is exponential in the lengths N and M. Therefore dynamic

programming was used to introduce an O(NM) algorithm that defines the prefix sequences $X(1:n) = (x1, ...xn)$ for $n[1:N]$ and $Y(1:m) = (y1, ...ym)$ for $m[1:M]$ and $setD(n,m) = DTW(X(1:n), Y(1:m))$.

4.4 Model Construction

The model is constructed in 3 steps as follows:

Step 1: for each feature f, apply the K-means clustering algorithm on the time warped vector space and create K clusters that span the n age bands. Note that we use K to distinguish it from the neighborhood parameter k used in the k-NN search. This will result in a total of $K * n$ centroids, one for each age band in each cluster.

Step 2: for each probe image its age band is determined. For that age band the cluster C_i that fits most closely to the image, given a fiducial feature X_j is identified on the basis of the conditional probability of the occurrence of the cluster given:

$$Pr(C_i|X_j) = Pr(C_i)Pr(X_j|C_i)/Pr(E) \tag{4.3}$$

where $Pr(E)$ is a normalizing factor taken across all clusters. In practice the normalizing constant $Pr(E)$ need not be computing as maximum likelihood is applied on (4.3) above thus resulting in retruning the clusters which maxmimizes $Pr(C_i|X_j)$.

Step 3: Use equation (3.2) with M_i representing the mean of feature i across the cluster that the image belongs to, instead of the mean across all images in the database. Thereafter, equation (3.3) is applied to find the pivot points P_i and k polynomial

aging function of the same form as in equation (3.4) is developed for each of the K clusters.

Algorithm 1 Local Model

Input: Probe images and its data for features (X_1....X_n).
Gallery images and features (X_1.....X_n) of each image.
Output: Rank determined.

```
for each probe image do
    Read its features (X1.....Xn) data.
    Get probability Pr of cluster (C1.....Ck) by using equation 4.3
    Sort the clusters (C1.....Ck) in descending order of probability.
    Choose top M clusters out of the K clusters.
    for each cluster (C1.....CM) do
        for each feature (X1.....Xn) do
            Use equation 3.2 and calculate deviation factor
            for each band(b1.....bn) do
                Use equation 3.3 and calculate de-aged distance
                Use the knn search algorithm and find the closest k neighbors
                 of de-aged distance
                Get centorid of k neighbors
            end for
            Use equation 3.4 and determine a quardratic polynomial function .
            Get predicted distance vector P of all features (X1.....Xn)
        end for
        for  each of M clusters do
            Calculate subspace S as per method explained in section 3.3
            Use predicted distance and determine similarity matrix.
            if rank1 is obtained then
                break
            end if
        end for
        Final rank is obtained
    end for
    Collate the results for the rank of all images.
end for
```

4.5 Empirical Study

The experiments were performed with the same FG-NET and MORPH databases as used in previous chapters. We have used the same settings for agebands and subspaces throught out this experimentation.

4.6 Experiment with FG-NET Database

For evaluation we chose all 1002 images taken from the entire set of 82 individuals. A leave-one person-out (LOPO) strategy was used for evaluation (Gong et al., 2013, 2015; Li et al., 2011). The same method is used for age subspace selection as given in the previous chapter. The algorithm for ranking images using the local modelling approch is given below in Algorithm 1.

4.6.1 Comparison of models across data granularity level for shape features

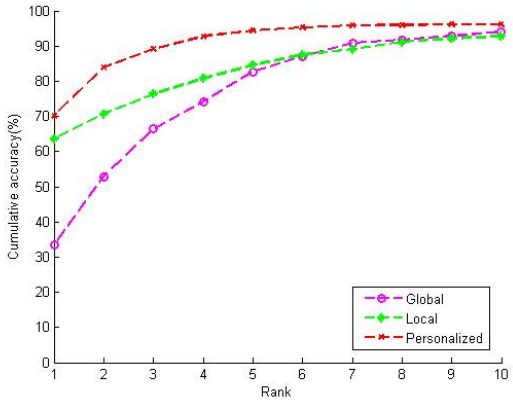


Figure 4.3: Cumulative Matching Characteristic Curves (CMC) for FG-NET Database

The CMC curves in Figure 4.2 confirm that the trends displayed in shape models hold across rank orders from 1 to 10, with the Personalized model emerging the winner across the range. However, it is apparent that the local modelling approach is superior to that of the global approach, especially at the lower rank orders.

4.6.2 Comparison of models across data granularity level for texture features

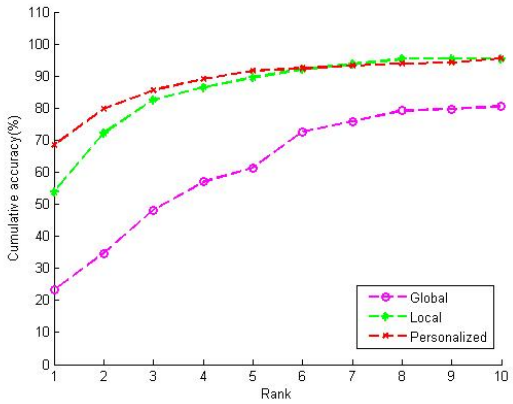


Figure 4.4: Cumulative Matching Characteristic Curves for FG-NET Database

The CMC curves in Figure 4.3 again confirm that the trends displayed for shape features also hold for the texture features as well, with the personalized approach emerging the winner once again followed by the local and global approaches.

Table 4.2 reveals the advantges of the local modelling approach over the personalized approach. While its rank 1 accuarcy (Table 4.3) is lower than that of the personalized approach its performance with respect to average probe image time is substantially less than that of the former approach.

Table 4.1: Comparative results of Rank-1 score with FG-NET database

Models	**shape**	**Texture**
Global	33.7%	23.3%
Personalized	70.4%	68.6%
Local	63.7%	53.9%

Table 4.2: Comparative results of Time per probe Image with FG-NET database

Models	**Shape**	**Texture**
Global	0.11s	0.9s
Personalized	0.58s	0.56s
Local	0.27s	0.24s

4.7 MORPH Experimentation

4.7.1 Comparison of models across data granularity level for shape features

We next tested the performance of the local modelling approach on the MORPH gallery. Given that this gallery differs markedly from FG-NET in terms of having a much smaller average age gap between two images from the same person it would be interesting to see the effects of this characteristic on rank 1 accuracy. The expectation would be that the local modelling approch would perform significantly better than its counterparts with respect to usage of both shape and texture features.

Table 4.3: Comparative results of Rank-1 score with FG-NET database

Models	**Rank-1**
Global (Texture)	23.3%
Global (Shape)	33.7%
Personalized(Texture)	68.6%
Personalized(Shape)	70.4%
Local(Texture)	53.9%
Local(Shape)	63.7%
(Geng et al., 2007)	38.1%
(Park et al., 2010)	37.4%
(Li et al., 2011)	47.5%
(Gong et al., 2013)	69%
(Gong et al., 2015)	76.2%

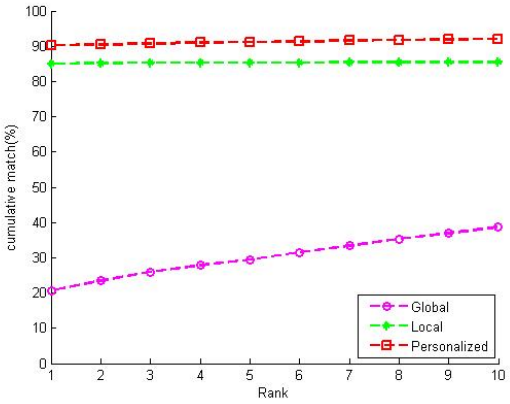


Figure 4.5: Cumulative Matching Characteristic Curves for MORPH Database

The CMC curves in Figure 4.5 confirm that the the Personalized model is once again the winner across rank orders 1 to 10. Howvever, we note that the local modelling approach has closed the gap with the Personalized model, in keeeping with our expectations.

4.7.2 Comparison of models across data granularity level for texture features

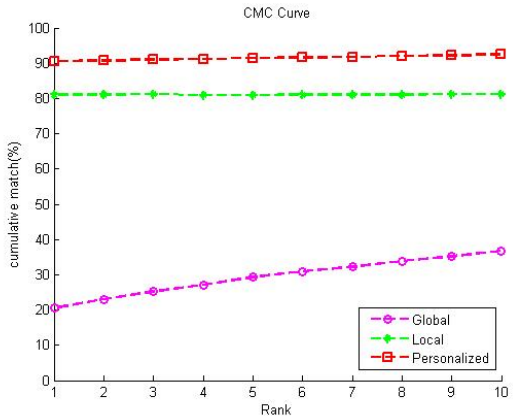


Figure 4.6: Cumulative Matching Characteristic Curves for MORPH Database

The CMC curves in Figure 4.6 reveals that the trends displayed for shape features also hold for texture features across rank orders from 1 to 10, with the Personalized model emerging the winner across the range, closely followed by the Local model.

Table 4.4: Comparative results of Rank-1 score with MORPH database

Models	**Shape**	**Texture**
Global	21%	20%
Personalized	91%	90.5%
Local	85%	81%

The results in Tables 4.1, 4.2, 4.4 and 4.5 clearly indicate that the Local approach for shape and texture returns good Rank-1 accuracy and relatively low recognition time as compared to Personalized models. However, Personalized models for shape

Table 4.5: Comparative results of Time per probe Image with MORPH database

Models	**Shape**	**Texture**
Global	0.18s	0.16s
Personalized	1.43s	1.36s
Local	0.56s	0.51s

Table 4.6: Comparative results of Rank-1 score with MORPH database

Models	**Rank-1**
Global(Texture)	20%
Global(Shape)	21%
Personalized(Texture)	90.5%
Personalized(Shape)	91%
Local(Texture)	81%
Local(Shape)	85%
(Klare & Jain, 2011)	79.08%
(Park et al., 2010)	79.80%
(Li et al., 2011)	83.90%
(Gong et al., 2013)	91.14%
(Gong et al., 2015)	94.59%
(Li et al., 2016)	94.87%

and texture do return better rank-1 accuracy as compared to Local and Global models. Tables 4.3 and 4.6 show that the Local model performance is also comparable to state of art approaches(Klare & Jain, 2011; Park et al., 2010; Li et al., 2011; Geng et al., 2007) for FG-NET and MORPH database. However, we note that rank1 accuracy is still short of what (Gong et al., 2015) achieved. In the next chapter we present new versions of our aging modelling scheme based on two new models, namely *Integrated* and *Composite* that perform competitively to (Gong et al., 2015).

4.8 Discussion

In this chapter, Local models for shape and texture are presented. In the case of the Local model we reduced the search space by first dividing the whole database into sub populations(clusters) and then considered subspaces within each subpopulation. The Local model trades-off accuracy with recognition time. Reduction in recognition time follows directly from restricting the search space to a cluster rather than the entire dataset. Experimentation with the FG-NET and MORPH database resulted in a reduction of recognition time in relation to the Personalized model approach. But a reduction in rank 1 accuracy also resulted due to the fact that image recognition decisions are made on localized data rather than on the entire dataset. After building the three models viz Global, Personalized and Local at different level of data granularity a question arises "Does any synergy exist between these three models? ".

In the next chapter we will investigate this question and search and utilize the synergy to build integrated model where we combine these three models.

Chapter 5

Integrated Modelling

5.1 Introduction

The characterics of features of the human face indicates that there are certain trends which are globally true for whole of the problem space. On the other hand, there are other trends that are common to only subsets of population and not valid for whole of the population. The former set of chatacteristics determines a global model and the latter set determines a local model in our study. There is yet a third set of trends that is characteristically special to the individual image. Such trends are studied under the broad concept of a personalized model. All these three approaches provide complementary information and data that is useful for inter related complex modelling tasks. It is also hypothesised that if a synergy exists between these approaches then an integrated model approach can be used by combining all the types of information into a single cohesive model.

(Kasabov, 2007b) has also introduced the concept of an integrated scheme embracing different types and levels of knowledge. The author discusses the prediction

capability of each of the three models (global, local and personalized) and then proposes a single multimodel system that encompasses all three models. However he did not give an implementation of his integrated framework. (Widiputra, 2011) introduced the integrated multi model framework for multiple time series. A methodology was laid down for construction of such a synergetic integrated model that deals with multiple time series along with their interactions and profiles of relationships. It utilizes information from all the three types of models.

The experiments and analysis carried out in this chapter will answer the questions: does any synergy exist in the three approaches (global, local and personalized) with a view to enhancing overall accuracy.

5.2 Integrated Model

Our Integrated model is inspired by the Integrated Multi Model Frame Work(IMMF) proposed by (Widiputra, 2011) which exploits synergy that may exist between models at the global, local and personalized data granularity levels.

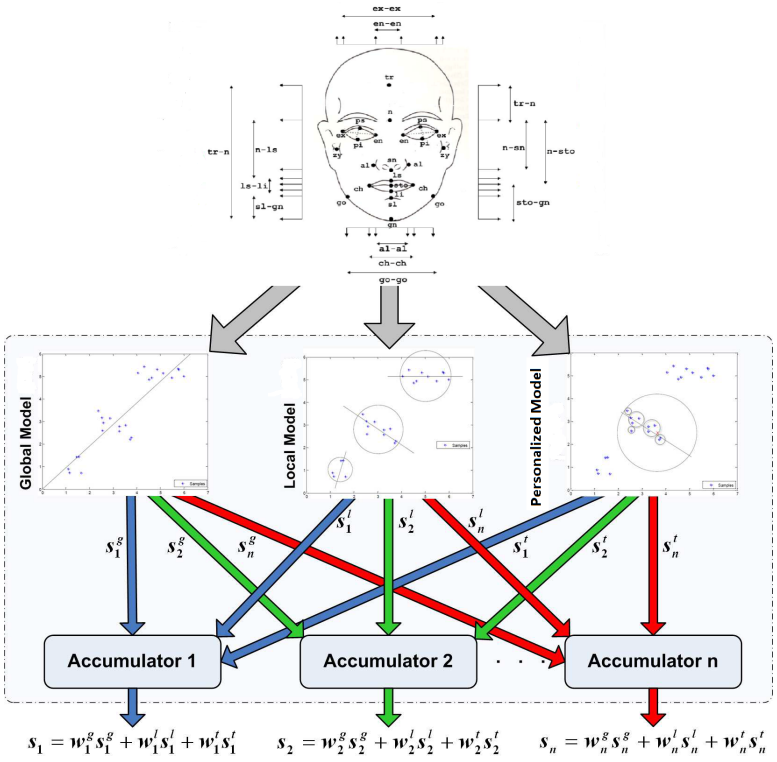


Figure 5.1: Illustration of integrated framework of global, local and personalized models. (Hwang,2009, Widiputra et al., 2011)

The main component of the IntegratedMulti-Model Framework (IMMF) is the accumulator module which is implemented through the use of an Adaline Neural Network (Widrow & Stearns, 1985) that decides the contribution (weighting) of each type of model. The accumulator module takes as input the set of training images and then for each training image will feed prediction outcomes (in the form of predicted feature value for feature j) from each of the global, local and personalized models in order to assign weights that will be associated with each type of model. The output of the accumulator represented by X in Figure. 5.2 is the optimized final set of weights over the entire set of training images. Equation 5.1 defines the prediction for feature

j for the Integrated model in terms of a linear weighted function of predictions from the global, local and personalized models.

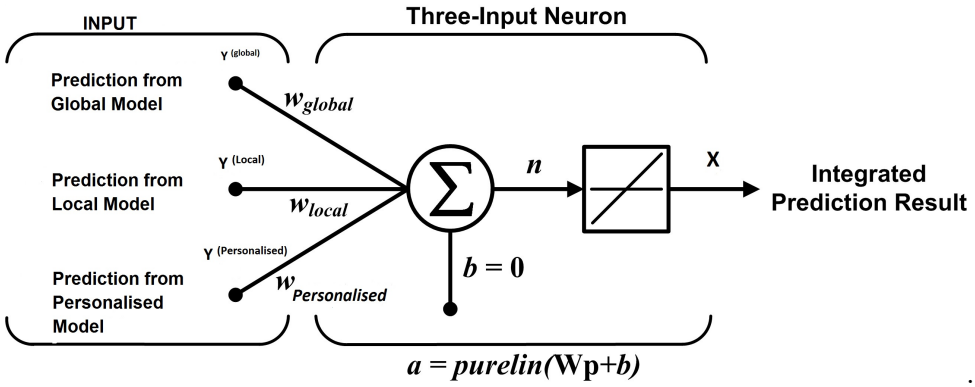


Figure 5.2: Illustration of the Adaline for Integrated Model

$$(X)_j = W_{j,g}Y^{(Global)} + W_{j,l}Y^{(Local)} + W_{j,p}Y^{(Personalised)} \tag{5.1}$$

where $W_{j,g}$ is the weight attributed to the global model for fiducial feature j. $W_{j,l}$ is the weight attributed to the local model for fiducial feature j. $W_{j,p}$ is the weight attributed to the personalized model for fiducial feature j. $Y^{(Global)}$, $Y^{(Local)}$ and $Y^{(Personalised)}$ are values returned from equation 3.1 and 3.4.

ADALINE Learning Rules

The ADALINE network adjusts the connection weight according to the weighted sum of the inputs, whilst in a different way, the standard perceptron adjusts the connection weights according to the output of the activation or transfer function. Nevertheless, this methodology considers that there is no external interference to the system. In relation to that, the bias value b in the ADALINE is not considered. Therefore, only the weights are being estimated through each iteration.

In general, the Least Mean Square (LMS) algorithm proposed by (Rappaport et al., 1996; Haykin, Haykin, Haykin, & Haykin, 2009) adjusts the weights and biases (if any is being considered) of the ADALINE network in order to minimise the mean square error, where the error is the difference between the target output and the network output.

The key insight of the learning rule was to estimate the ADALINE networks mean square error F(x) defined by:

$$F(x) = e^2(k) = (t(k) - a(k))^2, \tag{5.2}$$

where the error e is a function of the weights vector w. Consequently, as weights change, the error changes. The objective of the learning process is then to move in weight space down the slope of the error function with respect to each weight. Nevertheless, the size of the move should be proportional to the magnitude of the slope. Then at each iteration k a gradient estimate of the form that is given below can be obtained:

$$\nabla F(x) = \nabla e^2(k) \tag{5.3}$$

$$x_k + 1 = x_k \alpha \nabla F(x)|_{x=x_k} \tag{5.4}$$

$$= x_k + 2\alpha e(k) z(k) \tag{5.5}$$

The equations of the LMS algorithm can then be written conveniently in matrix notation:

$$w(k+1) = w(k) + 2\alpha e(k) p^T(k) \tag{5.6}$$

where the error e and the bias b are now vectors.

$$b(k+1) = b(k) + 2\alpha e(k) \tag{5.7}$$

Equations 5.6 and 5.7 make up the LMS algorithm which is the learning algorithm for the ADALINE network (also known as the Widrow-Hoff learning algorithm or the delta rule). However, since no bias value is being considered in the integrated framework proposed in this study, what is being estimated through each iteration are only the weights. Additionally, a single accumulator module in the IMMF would produce only a single output (as the integrated prediction results from the global, local and transductive model) and not multiple outputs. Therefore, in the learning algorithm of the IMMF only Equation 5.6 is used.

Learning Algorithm of Integrated Model

The learning procedure implements two stopping conditions of the ADALINE learning process when finding the optimal weights with the smallest mean squared error. The first stopping condition is the maximum number of epochs, while the second one is the tolerance level for the mean squared error.

Step 1: For each feature(X_1.....X_n) of training image , a predicted distance vector from global, local and personalized model is calculated for feature j. These values are multiplied with the weights vector as shown in equation 5.1. In each case the error is determined by comparing predicted distance with the actual feature value of image.

Step 2: The weights are revised by using equation 5.6 and the same process (step1) is repeated for all training images. When all the images are processed, a final weight vector is determined. These weights are used to obtain the mean squared error. If the squared error is less than or equal to the tolerance value, then the weights for that feature are finalized.

Step 3: If the error is greater than tolerance value then steps 1 and 2 are repeated. Each such iteration is called an epoch. This proccess is repeated for the maximum number of epochs. If error falls below the tolerance before the maximum number of epochs is reached then the process is stopped. If not, we continue for maximum number of epochs.

Step 4: Finally weight vector for the feature is determined. This process is repeated for

each feature and a weight vector finalized for each feature.
These weights are used for preparation of Integrated model and judging the overall performance of the system.

In the preparation of Integrated model, three parameters need to be chosen. These are learning rate α, number of epochs and the tolerance level. The other two are stopping criterion of the learning process for determining weights. During the conduct of experiments the learning rate is set to 0.1, while number of epochs is set to 200 and the tolerance level is set to $1e^{-5}$.

Algorithm 2 Integrated Model

Input: Probe images and its data for features $(X_1....X_n)$.
Weights calculated during training.
Gallery images and features $(X_1.....X_n)$ of each image.
Output: Rank determined.

1: **for** each probe image **do**
2: Read its features $(X_1.....X_n)$ data.
3: Get predicted distance for global model as explained in section 3.2.1
4: Get predicted distance for Personalized model as explained in section 3.2.2
5: Get predicted distance for local model as explained in section 4.4
6: Use equation 5.1 and get weighted output $(X)_j$ for features.
7: Calculate subspace S as per method explained in section 3.3
8: Use weighted output $(X)_j$ and determine similarity matrix.
9: Final rank is obtained
10: **end for**
11: Collate the results for the rank of all images.

5.3 Empirical Study

The experiments were performed with same FG-NET and MORPH databases as used in previous chapters. We have used the same settings for agebands and subspaces through out

this experimentation.

5.4 Experiment with FG-NET Database

For evaluation we chose all 1002 images taken from the entire set of 82 individuals. A leave-one person-out (LOPO) strategy was used for evaluation (Gong et al., 2013, 2015; Li et al., 2011). The same method is used for subspace selection as used in previous chapter.

5.4.1 Comparison of models across data granularity level for shape features

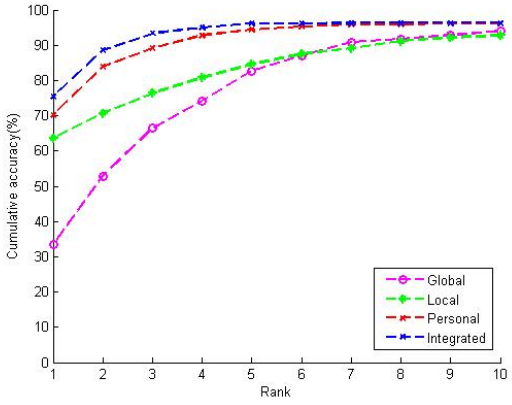


Figure 5.3: Cumulative Matching Characteristic Curves (CMC) for FG-NET Database

The CMC curves in Figure 5.3 confirm that the trends displayed for shape models hold across all rank orders from 1 to 10, with the Integrated model emerging the winner across the range.

5.4.2 Comparison of models across data granularity level for texture features

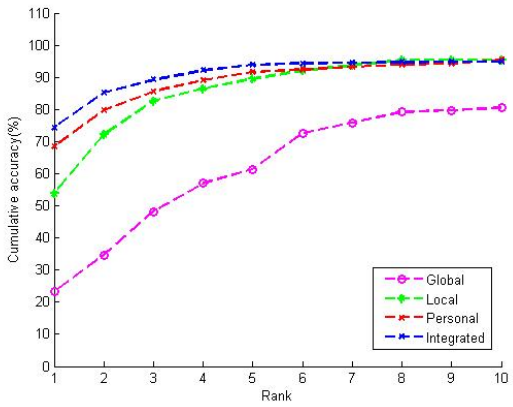


Figure 5.4: Cumulative Matching Characteristic Curves for FG-NET Database

The CMC curves in Figure 5.4 again confirm that the trends displayed for texture models holds across all rank orders from 1 to 10, with the Integrated model emerging the winner across the range.

Table 5.1: Comparative results of Rank-1 score with FG-NET database

Models	**shape**	**Texture**
Global	33.7%	23.3%
Personalized	70.4%	68.6%
Local	63.7%	53.9%
Integrated	75.6%	74.5%

Table 5.2: Comparative results of Time per probe Image with FG-NET database

Models	Shape	Texture
Global	0.11s	0.9s
Personalized	0.58s	0.56s
Local	0.27s	0.24s
Integrated	0.78s	0.75s

Table 5.3: Comparative results of Rank-1 score with FG-NET database

Models	Rank-1
Global (Texture)	23.3%
Global (Shape)	33.7%
Personalized(Texture)	68.6%
Personalized(Shape)	70.4%
Local(Texture)	53.9%
Local(Shape)	63.7%
Integrated(Texture)	74.5%
Integrated(Shape)	75.6%
(Geng et al., 2007)	38.1%
(Park et al., 2010)	37.4%
(Li et al., 2011)	47.5%
(Gong et al., 2013)	69%
(Gong et al., 2015)	76.2%

5.5 MORPH Experimentation

5.5.1 Comparison of models across data granularity level for shape features

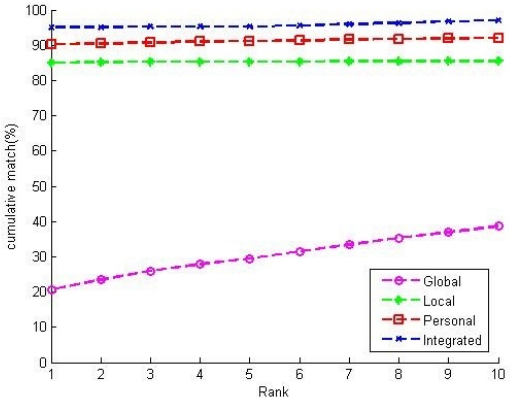


Figure 5.5: Cumulative Matching Characteristic Curves for MORPH Database

The CMC curves in Figure 5.5 confirm that the trends displayed for shape models holds across all rank orders from 1 to 10, with the Integrated model emerging the winner across the range.

5.5.2 Comparison of models across data granularity level for texture features

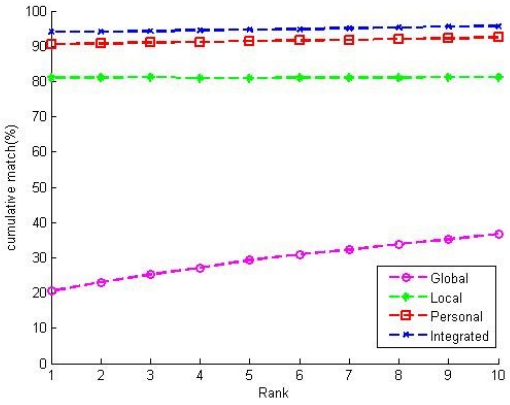


Figure 5.6: Cumulative Matching Characteristic Curves for MORPH Database

The CMC curves in Figure 5.6 again confirm that the trends displayed for texture models holds across all rank orders from 1 to 10, with the Integrated model emerging the winner across the range.

The results in table 5.1, 5.2, 5.4 and 5.5 clearly indicate that Integrated approach for shape and texture returns the highest Rank-1 accuracy at the cost of more processing time. Tables 5.3 and 5.6 show that the Integrated model performance is also competitive when compared to the state of art algorithms(Klare & Jain, 2011; Park et al., 2010; Li et al., 2011; Geng et al., 2007) for FG-NET and MORPH database. Our experimentation results accuracy surpasses that of (Gong et al., 2015) with MORPH database but falls slightly

Table 5.4: Comparative results of Rank-1 score with MORPH database

Models	**Shape**	**Texture**
Global	21%	20%
Personalized	91%	90.5%
Local	85%	81%
Integrated	95.3%	95%

Table 5.5: Comparative results of Time per probe Image with MORPH database

Models	**Shape**	**Texture**
Global	0.18s	0.16s
Personalized	1.43s	1.36s
Local	0.56s	0.51s
Integrated	1.84s	1.88s

short in the case of the FG-NET database. In the next chapter we present our Composite model that is competitive to that of (Gong et al., 2015) for the FG-NET database as well.

5.6 Discussion

In this chapter an Integrated model, generated separately for shape and texture features, is presented. The Integrated model combines three models viz Global, Local and Personalized models. It gives higher accuracy than all of other models on their own but at the cost of rise in computation time. Experimentation with the FG-NET and MORPH databases both corroborate this fact. The results show that there synergy exists between the three models that at different levels of data granularity. It was further observed that the Integrated model for shape and texture features sometimes retrun different matched images (to a given proble image) in their respective Rank 1 sets. This observation leads to further research where

Table 5.6: Comparative results of Rank-1 score with MORPH database

Models	**Rank-1**
Global(Texture)	20%
Global(Shape)	21%
Personalized(Texture)	90.5%
Personalized(Shape)	91%
Local(Texture)	81%
Local(Shape)	85%
Integrated(Texture)	95%
Integrated(Shape)	95.3%
(Klare & Jain, 2011)	79.08%
(Park et al., 2010)	79.80%
(Li et al., 2011)	83.90%
(Gong et al., 2013)	91.14%
(Gong et al., 2015)	94.59%
(Li et al., 2016)	94.87%

we introduced a method to remove the conflict when shape and texture do not both allot the same image to their Rank 1 sets. For this case we introduced a composite model which resolves the conflict between the Integrated models for shape and texture.

Chapter 6

Composite Modelling

6.1 Introduction

In this Chapter we present an approach whereby shape and texture models are combined in order to exploit synergy that exists between them. In certain cases, both types of models point to the same image which also happens to be the right one. In such cases having two types of models enhances the robustness of the matching process as agreement exists. However in certain other cases, each model could identify different images. Conflict resolution is then necessary and we make use of a decision tree classifier for resolving such conflicts. Our choice of the decision tree classifier was based on the fact that it had the highest conflict resolution success rate out of the classifiers that we experimented with, which included the Support Vector Machine (SVM) (Furey et al., 2000), Decision Tree (Schmid, 2013; Kuncheva, 2004; Friedl & Brodley, 1997), Random Forest (Liaw & Wiener, 2002) and Naive Bayes(Amor, Benferhat, & Elouedi, 2004).

We built separate models for shape (four models, at diffent levels of data granularity) and texture (four models) and obtained separate results for each. A composite model was then built after resolving conflicts between shape and texture models in order to make a final decision on the probe image. In certain cases, both models (shape and texture) point to

same person which is also the right one. But in other cases, each of the two models predicts a different person. In the conflict situation we used a decision tree classifier for resolving the conflict. Our empirical study show that the composite approach further improves rank 1 accuracy over the Integrated approach that we presented earlier in Chapter 5.

6.2 Composite Model

There are two types of scenarios possible with both types of models operating in tandem. The first is that both model types return the same predicted image for a given probe image. This scenario increases confidence in the prediction process, and thus in this case the returned image is selected. The other scenario occurs when both types of models point to different images. Figure 6.1 illustrates the Composite model approach.

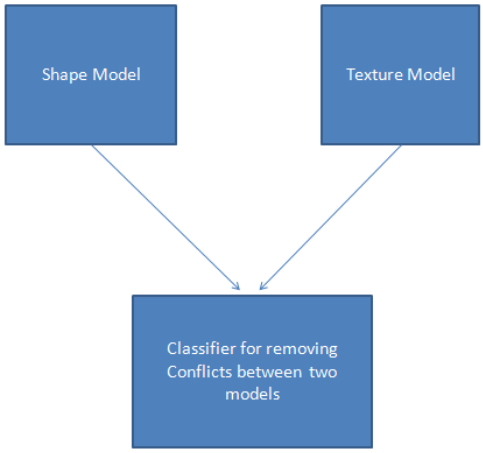


Figure 6.1: Composite Model

The conflict scenario is a non trivial case to deal with as it requires a conflict resolution mechanism. The resolution mechanism that we adopt is a classifier. The conflict scenario can be represented as a two class classification problem: class 1 corresponds to the case when

the belief is that the shape model type has made the right decision while class 2 corresponds to the belief that the texture model has made the right decision. In principle, any type of robust classifier can be utilized here, including the Bayesian that we used for subspace identification. We experimented with different classifiers such as the support vector machine (SVM), Decision tree, Random Forest and Naive Bayes, and found that the Decision Tree returned the highest accuracy amongst all.

The predictors used were: the 7 shape features, 5 texture features and 2 deviation factors df_s and df_t; one each for shape and texture respectively. In addition, we use two confidence measures c_s and c_t, for shape and texture respectively, which we define below.

The confidence factor is a measure of the predictive power of the model type on identifying the correct image in the gallery. The closer the distance between the predicted image returned by the model to the desired image, the larger should be the confidence. However, it is desirable that the confidence diminishes rapidly in a non-linear manner with distance; a distance of $2d$ between the images should be considered much more than twice as weak as a confidence obtained with a distance of d. Hence our decision to choose an inverse exponential function to measure confidence. In principle, any monotonically decreasing nonlinear function can be used, but our experimentation with different types of such functions has shown that the inverse exponential function defined in equation 6.1 performs best.

$$conf_m = 1/(1 + e_m^d) \tag{6.1}$$

where subscript m refers to the model type, taking values of either "shape" or "texture" and e is the natural logarithm base

For either of the model types, the distance can be represented as a Euclidean distance between the predictor image feature vector and the actual image feature vector and is defined in expression 6.2 below.

$$E(P, A) = \sqrt{\sum_i (P_i - A_i)^2} \tag{6.2}$$

The confidence measures for both model types were generated from the gallery dataset.

Algorithm 3 Composite Model

Input: Results of testing the shape and texture model.
Input: feature($X_1, X_2, ...X_n$) values of shape and texture
Output: Final Rank determined.

1: **for** each probe image **do**
2: **if** (Shape Rank==1 && Texture Rank $\sim$ =1) ||(Shape Rank $\sim$=1 && Texture Rank== 1) **then**
Get features $(X_1, X_2, ...X_n)$ from Shape Model
Get feaures $(X_1, X_2, ...X_n)$ from Texture Model
Get deviation factor by using equation 3.1
Calculate Confidence measure for Shape and Texture by using equation 6.1
Prepare feature vector $V = [X_1.....Xn, C_s, C_t, df_s, df_t]$
3: **end if**
Use a classifier as described in section 6.2 and obtain classification.
Final rank is obtained from the classification
4: **end for**
Collate the results for the rank of all images.

6.3 Empirical Study

The experiments were performed with the same FG-NET and MORPH databases as used in previous chapters. We have used the same settings through out the experimentation.

6.4 Integration vs Composition

In this set of experiments we investigated the effects of combining shape and texture models at different levels of granularity. We benchmarked our models against the current state-of-the-art one proposed by (Gong et al., 2015).

Table 6.1: Comparative results of Rank-1 score with FG-NET database

Models	**Rank-1**
Global	41.7%
Local	65.1%
Personalized	71.2%
Integrated	77.6%
(Geng et al., 2007)	38.1%
(Park et al., 2010)	37.4%
(Li et al., 2011)	47.5%
(Gong et al., 2013)	69%
(Gong et al., 2015)	76.2%

The results in Tables 6.1 and 6.2 clearly indicate that the Composite Local and Personalized approaches significantly outperform the Global approach. In fact, the Composite Local and Personalized models are capable of good accuracy by themselves. This is to be expected as aging is a personalized process and its trajectory is very different for different individuals, depending on a range of factors such as lifestyle, genetic disposition and others. At the same time we observe that the Composite Integrated approach provides further improvement to models operating on their own, indicating that a certain amount of synergy existed between the two modeling approaches. We also note that the Composite Integrated approach significantly outperformed the models proposed by (Gong et al., 2015) and (Li et al., 2016). The Integrated vs Composite results are paired for each test image and a paired t-test is used for testing. Our composite model results are statistically significant with a P value for FG-NET of 0.00001 and a P value of 0.0000001 for the MORPH database based on a confidence level of 95%.

Table 6.2: Comparative results of Rank-1 score with MORPH database

Models	**Rank-1**
Global	21.2%
Local	81.2%
Personalized	90.6%
Integrated	96.7%
(Klare & Jain, 2011)	79.08%
(Park et al., 2010)	79.80%
(Li et al., 2011)	83.90%
(Gong et al., 2013)	91.14%
(Gong et al., 2015)	94.59%
(Li et al., 2016)	94.87%

6.5 Discussion

Based on the results of our Empirical study, we observe that model composition, similar to model integration, resulted in better accuracy. The final conclusion of the study was that the optimal modeling approach is to compose shape models built by integration across data granularity levels with their corresponding texture counterparts using the same data granularity scheme.

A brief comparison in tabular form is given in Table 6.3. Table 6.4 compares the models on two factors, complexity and effectiveness. Complexity relates to computation time and effort involved in model generation, whereas effectiveness relates to accuracy. Tables 6.3 and 6.4 provides some guidelines asto which models should be used, depending on the end-user's requirements. For highest accuracy the Integrated Model is suggested; on the other hand when time is at a premium, the Local Model is a good compromise as it has a lower recognition time than the Personalized model while having better accuracy than the Global model.

Our aging models are also capable of forward traversal through time. Forward traversal

Table 6.3: Complexity and Effectiveness Characteristics of Models

Models	**Complexity**	**Effectiveness**
Global	Low	Low
Local	Low	High
Personalized	High	High
Integrated	High	High

Table 6.4: Usage Criteria for Aging Models

Models	**When to Used**
Global	When results are required in a very short time
Local	When good accuracy is required in short time
Personalized	When high accuracy is required and recognition time is not a problem
Integrated	When very high accuracy is required and recognition time is not a problem

has two attractive benefits. The first is automatic update of the image database. Over a period of time an image gallery containing older images of a person becomes less useful for recognition purposes as research has shown that the longer the time gap between the probe image and the target image, the lower is the rank 1 accuracy in general (Park et al., 2010). The solution to this problem either involves obtaining newer images from the subjects concerned which may not be practicable as some subjects may not be available. A more practical alternative would be an automatic refresh based on a computed trajectory provided by the aging functions proposed in this research. The second benefit lies in age simulation whereby the appearance of given person at a future point in time is required for

applications as face animation or for making decisions at a future point in time on plastic surgery alternatives.

In next chapter we will discuss extraction of knowledge about aging process and show simulation process with our aging models.

Chapter 7

Knowledge Discovery

7.1 Introduction

Aging is a complex process because different changes would occur in a face at different points on the aging trajectory. The aging process could affect craniofacial growth, texture, soft skin changes, wrinkles and skin colour at different age points. It would be interesting to identify and search changes in the face at different age points. Our aging models can provide important insights into craniofacial growth of the face and how features change at different age points. This aging information can be used in forensic sciences(İşcan & Helmer, 1993; Neave, 1998), face animation and face recognition(Gates, 2011) technologies and in plastic surgery. It is important to understand that patterns, characteristics and rates of aging depend upon culture and lifestyle choices. Furthermore, environmental factors could play a role as well. A person's Gender could also play a role when taken in conjunction with lifestyle and environmental conditions. Finally, factors such as genetic makeup of the individual could also play an important part in the aging process. Thus, it evident that aging is influenced by a variety of factors and the complex interplay between these factors ultimately determines a given person's aging trajectory.

In order to fully account for all of these factors a rich database consisting of lifestyle variable and genome information needs to be available. Such databases are currently not available and the approach taken in this research has to use the limited information present in currently available databases to model craniofacial change with time.

7.2 Craniofacial growth

Aging is a personalized process and so we use our personalized aging models to build a simulated shape profile of a person across time. Firstly, we build a personalized aging model for any given person by applying the methods presented in section 3.3.2. For constructing a simulated shape profile we use equation 7.1 given below.

$$C_{updated} = C_i + (P_i - P_{i-1}) \tag{7.1}$$

In equation 7.1 the updated coordinates $C_{updated}$ of the face are obtained by adding $(P_i - P_{i-1})$ into coordinates C_i of the face, where P_i and P_{i-1} are predicted features value of two consecutive age bands and i denotes the age band index. For forward projection we build the aging model starting from age band 1 and move in the forward direction by using equations 3.2 and 3.3. To streamline the presentation, we reproduce equations equations 3.2 to 3.4 which are used in the algorithm below

$$DF_i = M_i - F_i \tag{3.2}$$

$$P_{ij} = G_{ij} + DF_i \tag{3.3}$$

$$y_i^{(personalized)} = c_1 {x_i}^n + c_2 {x_i}^{n-1} + ... + c_n x_i + c_{n+1} \tag{3.4}$$

Table 7.1: FG-NET Age Bands

Ages	0-5	6-10	11-15	16-20	21-25	26-30	31-35	36-40	41-45	46-69
Images	233	178	164	155	81	62	38	31	26	34
Subjects	75	70	71	68	46	38	30	24	19	10

Algorithm 4 Craniofacial growth

Input: Images and its data for features $(X_1....X_n)$.
Gallery images and features $(X_1.....X_n)$ of each image and 68 fiducial points.
Output: Predicted image.

1: **for** each probe image **do**
2: Read its features $(X_1.....X_n)$ data.
3: **for** each feature $(X_1.....X_n)$ **do**
4: Use equation 3.2 and calculate deviation factor
5: **for** each band$(b_1.....b_m)$ **do**
6: Use equation 3.3 and calculate de-aged distance
7: Use the knn search algorithm and find k neighbors of de-aged distance
8: Get centorid of k neighbors
9: **end for**
10: Use equation 3.4 and determine a quardratic polynomial function .
11: Get predicted distance vector P of all features $(X_1.....X_n)$
12: Update fiducial points by using equation 7.1
13: **end for**
14: Plot predicted fiducial points to obtain simulated image
15: **end for**

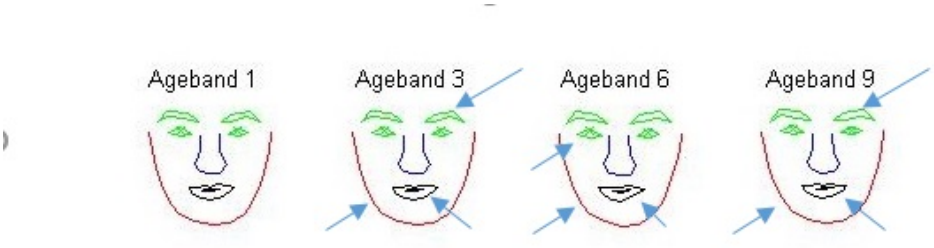
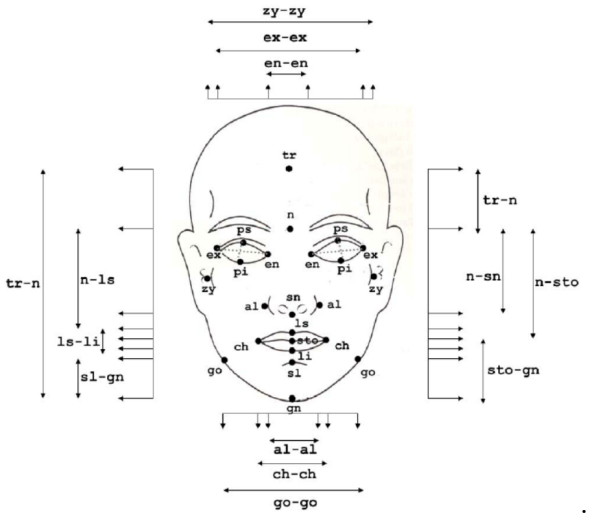


Figure 7.1: Personalized Craniofacial growth of a person at different age points (Arrows show parts where change occurs)

Personalised craniofacial growth of a person is shown in Figure 7.1. We can clearly see that as this person ages his fiducial features also change. By carefully examining the sketches we observe that craniofacial growth of the face occurs in horizontal and vertical directions until ageband 3. His mouth bone structure keeps changing till the ageband 9. There are little changes in eyes as compared to other features. We can also see changes in eyebrow bones at ageband 9. This could be the reason for wrinkles on the forehead and areas around the eyes. As discussed earlier aging is a personalized process and every person ages in different way. For general analysis we need to analyse the whole population. Therefore, in next the section we would discuss effects of anthropometric features at various points on the aging trajectory.

7.3 Effects of aging on shape



- Facial index ($\frac{n-gn}{zy-zy}$)
- Mandibular index ($\frac{sto-gn}{go-go}$)
- Intercantal index ($\frac{en-en}{ex-ex}$)
- Orbital width index ($\frac{ex-en}{en-en}$)
- Eye fissure index ($\frac{ps-pi}{ex-en}$)
- Vermilion height index ($\frac{ls-sto}{sto-li}$)
- Mouth Face width index ($\frac{ch-ch}{zy-zy}$)

Figure 7.2: Anthropometric Model and Seven Features.

After carefully examining Figures 7.2 and 7.3 we can extract information regarding the aging process in craniofacial features. Facial index is one such feature. It is defined as the ratio between ($\frac{n-gn}{zy-zy}$). It refers to the whole face. In this feature there is an increase in the ratio till the ageband 6. Beyond that point we observe a decrease. Mandibular index ($\frac{sto-gn}{go-go}$) refers to the height and width of the chin. This feature changes for the first three decades of the lifespan. After ageband 6 we observe a decrease in the value of this feature. Intercantal index ($\frac{en-en}{ex-ex}$) refers to the inner and outer width of eyes. There is a decrease in feature value till the fourth decade and after that we observe an increase in the feature value. Orbital width index ($\frac{ex-en}{en-en}$) refers to the eyes. Like intercantal index in the first four decades there is a slight increase in the feature value and after that it monotonically decreases. Vermilion height index ($\frac{ls-sto}{sto-li}$) refers to the lips. This feature decreases till

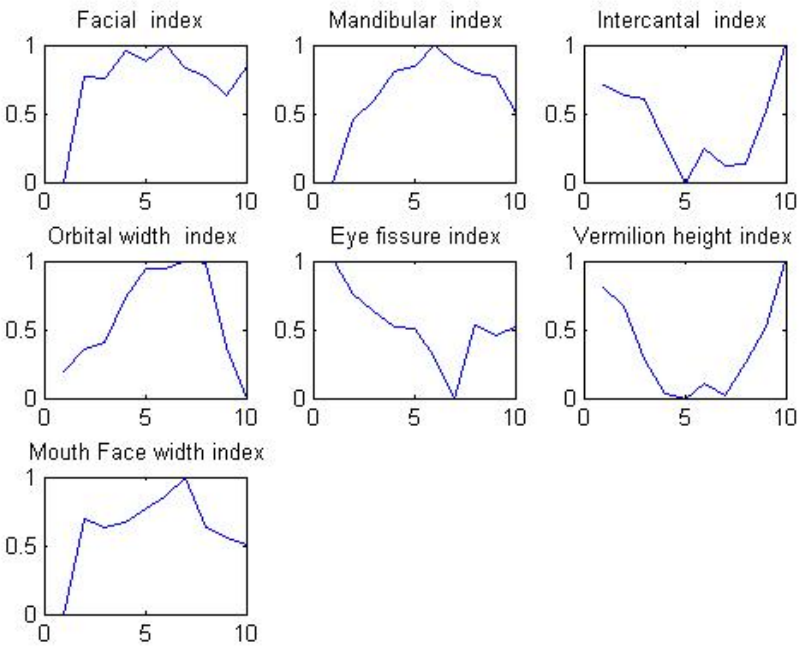


Figure 7.3: Shape features and their average behaviour at personalised level. Horizontal axis shows agebands and vertical axis shows feature values.

third decade and after that it would start increasing. Mouth Face width index ($\frac{ch-ch}{zy-zy}$) refers to the face width. For four decades there is an increase in this feature value and after that it appears to slightly decrease. The features that we constructed with horizontal measurements (en-en, ex-ex, go-go, zy-zy, ch-ch) and 3 vertical measurements (n-gn,n-sn, sto-gn)) are most representative of facial features. The features constructed with horizontal measurements increase until the fourth decade. We have used (Analysis of Variance) Anova with a 95 % confidence interval for verifying that the aging process is different between different age groups. Our results show that there is a significant change in features across the aging process.

Effects of aging on genders

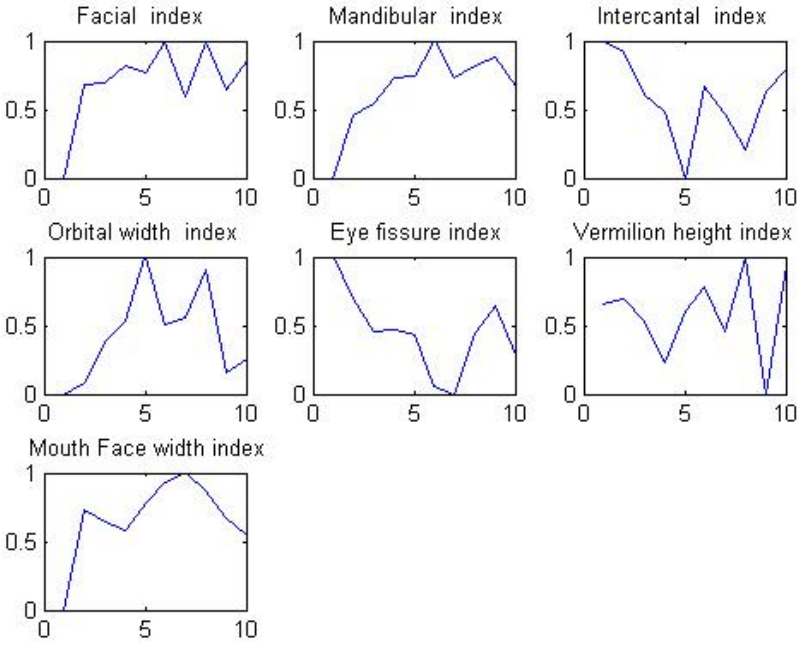


Figure 7.4: Female Shape features and their average behaviour at personalised level. Horizontal axis shows agebands and vertical axis shows feature values.

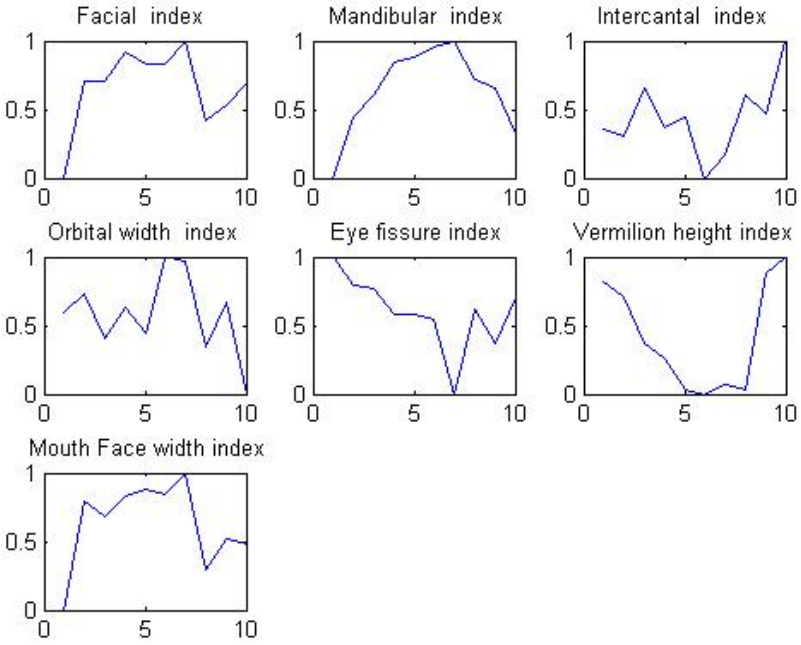


Figure 7.5: Male Shape features and their average behaviour at personalised level. Horizontal axis shows agebands and vertical axis shows feature values.

After carefully examining Figures 7.4 and 7.5 we can extract information regarding the aging process in craniofacial features between different genders. Facial index and Mouth Face width index aging patterns are similar but Mandibular index and Intercantal index, Orbital width index and Vermilion height index display different age trajectories across the two genders.

7.4 Effects of aging on texture

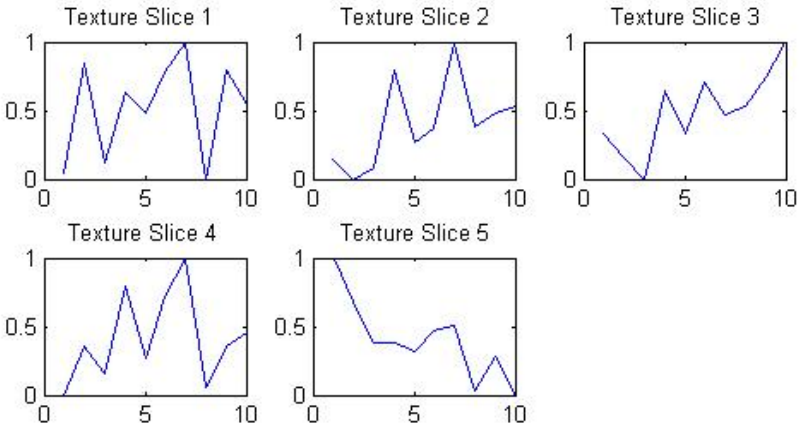


Figure 7.6: Texture features and their average behaviour at personalised level. Horizontal axis shows agebands and vertical axis shows feature values.

Figure 7.7: Five slices of the face for texture analysis.

In order to gain an understanding of the effect of texture features we divide the facial image horizontally into five slices. The slices cover the forehead, eyes, nose, mouth and the area below the mouth. Edges are marked on the face after converting the image into grayscale. A histogram is prepared for these edges. The frequency of the edges in each slice is taken as a feature, thus resulting in five features across the entire face as shown in Figure 7.6. We have used Anova with a 95% confidence interval and verified whether the aging process is different at different age points. Our results show there is significant change in features throughout the aging process. The number of edges on the face tends to increase as the person ages. This change can be detected by using Figures 7.6 and 7.7 for each texture slice. Texture slice 1 represents the forehead area. During aging the forehead feature shows that more edges form after ageband 6 and ageband 10. Texture slice 2 represents the eye area. During aging, the eye feature shows that more edges tend to form at ageband of 4 and 10. Texture slice 3 represents the nose area. It shows that during aging changes occur in the nose area throughout the lifespan of a person. Texture slice 4 represents the mouth area. It shows that during aging changes occur in the mouth at ageband of 4, 6 and 10. Texture slice 5 represents the chin area. It shows that during aging changes occur in the chin lesser occurs to a lesser extent when compared to other parts of the face. These changes could be more visible due to change in bone structure in the chin area.

Effects of gender on aging

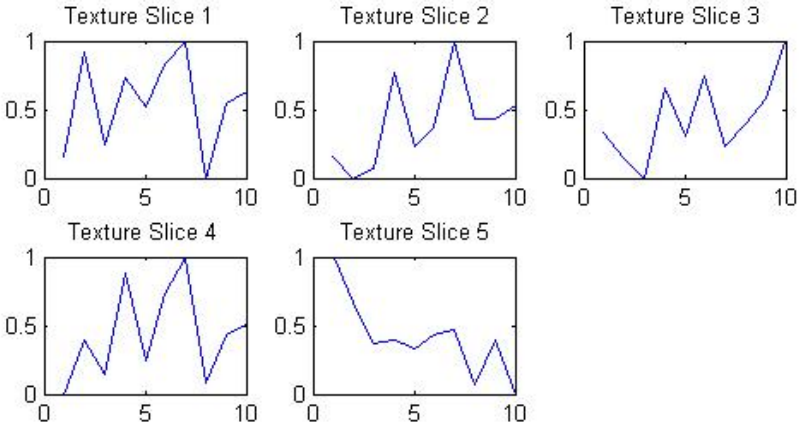


Figure 7.8: Female texture features and their average behaviour at personalised level.Horizontal axis shows agebands and vertical axis shows feature values.

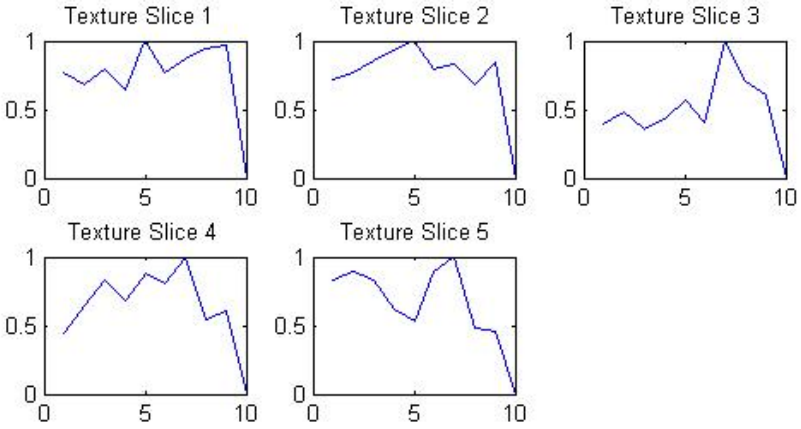


Figure 7.9: Male texture features and their average behaviour at personalised level. Horizontal axis shows agebands and vertical axis shows feature values.

After examining Figures 7.8 and 7.9 we can clearly see that texture changes in the two genders occur in different ways. Figure 7.8 for Females shows that the number of edges show an increasing trend with time. This implies that females would have more edges on the face at later ages such as ageband 10. On the other hand, Figure 7.9 for Males shows a downward trend in the number of edges. It means, at ageband 10 male would get lesser edges on the face than their female counterparts. This suggests that females age faster than males.

We validated the accuracy of our craniofacial growth simulation algorithm with the RMSE measure. The FG-NET database was used as the testbed for this purpose. The prediction error rate, as measured by the Root Mean Square Error (RMSE) is:

$$RMSE = \sqrt{\frac{\sum_i^n (\hat{P}_i - P_i)^2}{n}} \tag{7.2}$$

where $\hat{P}$, P represents predicted and actual coordinates respectively. The simulated model shows an RMSE value of 1.30 and 1.95 for (x,y) coordinates respectively. We also carried out a comparison of our algorithm with (Scandrett, Solomon, & Gibson, 2006). In order to standardize the comparison with the state of the art results reported from (Scandrett et al., 2006) , we predicted the average of the X and Y coordinates and obtained the RMSE of the average value. Table 7.1 shows clearly that our simulation method is significantly more accurate than the current state of the art method reported in the literature.

Table 7.2: Comparative results of RMSE with other methods

Models	**RMSE**
(Scandrett et al., 2006)	3.8
Proposed algorithm	1.6

7.5 Discussion

In this chapter, we have shown that our aging models are also capable of extracting useful knowledge through a forward traversal in time. We have discussed the aging simulation process that can be used for prediction of features at future time points. The knowledge extraction from aging models shows that there is significant change in fiducial features throughout the aging process. During aging vertical shape changes occur until around the ageband 4 (age of 20) and horizontal shape changes occur until the fourth decade of a person's life. The texture analysis also showed that change in features occur throughout life. These changes can be more visible at times because of a personalised life style. The craniofacial development of face and texture both play important role in aging process. It can be concluded that aging is a mixture of both texture and shape information (Park et al., 2010).

In the next chapter we will research gender recognition and age group classification.

Chapter 8

NeuCube Aging Model

8.1 Introduction

Aging is a slow process and its effects are visible only after a few years. But in spite of being slow, it remains a spatio-temporal phenomenon. The facial features of a person can be considered as a subspace and the aging over the years of this subspace is in turn a temporal process. It would be very useful to incorporate temporal as well as spatial patterns in aging data as important components in classification. Age group classification and gender recognition have important applications for business managers and law enforcement agencies. In Human Computer Interaction (HCI) gender recognition can be used to make interaction with humans more amenable to both genders. For example, it enables a computer to address a user by their correct title, Mr or Mrs, as the case may be. Automatic gender recognition facilitates better interaction with humans as well as saves keystrokes in filling up forms. In Surveillance systems, a gender specific physical locality can be concentrated upon, thereby reducing the area under observation and making the whole system more efficient. It could also assign higher threat levels to a specified gender location. In content based systems, indexing and searching can be greatly facilitated. For example, today a plethora of digital videos and photographs are produced. Gender based indexing and searching can be easily

carried out on this vast collection of digitised images. It will also reduce the search effort by limiting it to a particular gender, thereby making it more efficient. The same applies to Biometrics. The automatic collection of demographic data for statistical purposes will be facilitated if gender can be automatically recognised. Age group classification and gender recognition can play important role in age invariant face recognition. We use aging expert for finding the correct age band for model building in chapter 3. The age classification method can be incorporated with age invariant face recognition system for finding the correct age group. Moreover, gender recognition helps us build gender specific aging model for age invariant face recognition. This would help to improve the overall accuracy of age invariant face recognition system.

Age group classification and gender recognition are areas of research that is receiving a lot of attention due to advances in the field of biometrics (G. Guo & Mu, 2011; A. K. Jain et al., 2016; Liu, Yan, & Kuo, 2015; Ramanathan & Chellappa, 2006).

In this chapter we introduce a novel framework for age group classification and gender recognition. We use the same Anthropometric model for feature extraction and extract seven different indexes as in previous chapters. After extracting features we convert the data into a discrete spike train. We use the Address Event Representation (AER) encoding method to discretize the continuous signal. This encoding method was applied successfully for the artificial retina sensor (Hechenbichler & Schliep, 2004). We initialize the NeuCube by using the Small World rule (Hechenbichler & Schliep, 2004). At unsupervised training stage we used the STDP rule (Hechenbichler & Schliep, 2004). The dynamic evolving Spike Neural Networks (deSNN) is used here as an output classifier. Our choice of the NeuCube platform was influenced by its outstanding success that it has achieved on spatio-temporal classification problems in many diverse application domains (N. Kasabov et al., 2016; Dhoble, Nuntalid, Indiveri, & Kasabov, 2012; N. Kasabov, Dhoble, Nuntalid, & Indiveri, 2013; N. K. Kasabov, 2014). All such studies employing NeuCube have been on

data that exhibit a relatively fast pace of temporal changes. However in the aging domain that this research is based on, the speed of changes are much smaller, on the order of months to years rather than hours to seconds, and hence it will be interesting to investigate whether the success of NeuCube can be replicated in such a slow-changing temporal environment. Our empirical results indicate that NeuCube captures temporal patterns accurately in this domain as well, as its classification accuracy was significantly better than with classical methods such as the K Nearest Neighbor(KNN), Multi Layer Perceptron (MLP) and Naive Bayes classifiers.

8.2 Methodology

We have used the same anthropometric model with its seven features as discussed in the previous chapter. These features were fed into NeuCube. NeuCube, featuring a spiking neural network architecture that was successfully used in spatio-temporal modelling in brain data applications (N. K. Kasabov, 2014; N. Kasabov et al., 2014; N. Kasabov, 2012; N. Kasabov et al., 2015; Tu et al., 2014). In contrast to brain data signals obtained through the use of devices such as EEG where time length between observations spans seconds, the corresponding time length in facial aging spans years, thus making it an interesting experiment to test the efficacy of NeuCube's spatio-temporal modelling capability in an environment of much coarser time granularity. A block diagram of NeuCube's architecture is shown in Figure. 8.1. NeuCube consists of three major components: an input encoding module, a three-dimensional Spiking Neural Network Cube(SNNc), and an output dynamic evolving spiking neural network (deSNN) classifier (N. Kasabov, 2012). In the first stage we adjust the connection weights in spiking SNNc to learn spatio temporal patterns embedded in the data. In the second stage, supervised learning is carried out.

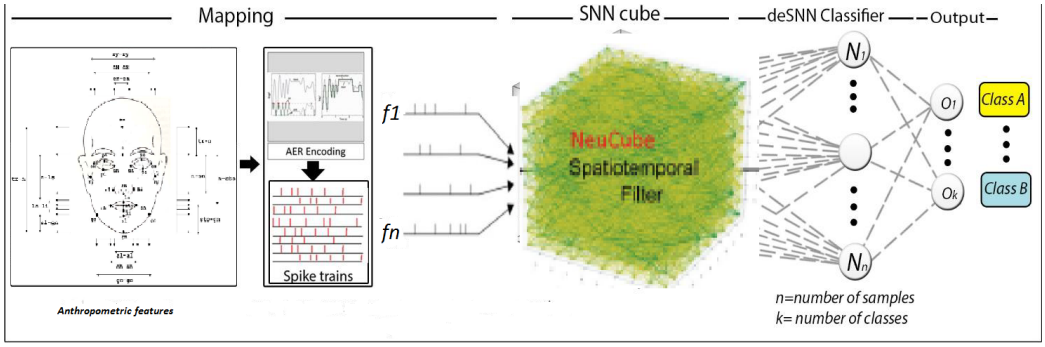


Figure 8.1: Schematic representation of the NeuCube-based methodology for mapping, learning, visualisation and classification of FG-NET DATA

Overall, the modeling process consists of data encoding, SNNc initialization, unsupervised training of the SNNc and finally supervised training of the classifier.

8.2.1 Data Encoding

In the aging problem feature data is in the form of real values, and hence needs to be discretised into spike trains. The Address Event Representation (AER) encoding method is used for this purpose. This encoding method was applied successfully for the artificial retina sensor(Hechenbichler & Schliep, 2004). The improved spike encoding method (SF) to recover the original signal is defined as for a given signal $S(t)$ where $(t = 1; 2; ::::; n)$, a baseline $B(t)$ variation during time t with $B(1) = S(1)$. If the incoming signal intensity $S(t_1)$ exceeds the base-line $B(t_{1-1})$ plus a threshold defined as Th, then a positive spike is encoded at time t_1, and $B(t_1)$ is updated as $B(t_1) = B(t_{1-1}) + Th$; and if $S(t_1) <= B(t_{1-1}) + Th$, a negative spike is generated and $B(t_1)$ is assigned as $B(t_1) = B_{t-1} + Th$. In other situations, no spike is generated and $B(t_1) = B(t_{1-1})$. We obtain both positive and negative spike trains for encoding. Positive spikes accompany an increase in feature value, whereas a negative spike occurs when the feature value has decreased(N. Kasabov et al., 2015).

8.2.2 NeuCube(ST) Initialization

We initialize the SNNc by the following rule. Each neuron in the SNNc is connected to its nearby neurons which are within a distance d, where d equals the longest distance between any pair of neurons in the SNNc multiplied by a parameter r. The initial weights are set with r selected at random in the range [-1, 1]. We randomly select 80% of the connection weights to be positive and the remaining 20% to be negative in the cube initialization process. The actual proportion of positive to negative connections will change continuously due to weight propagation that occurs when data is processed by the cube.

8.2.3 Training Stage I:

This stage deals with Unsupervised Reservoir Training. It encodes hidden spatio-temporal relationships from the input data into neuronal connection weights. According to the Hebbian learning rule, if the interaction between two neurons persists, then the connection between them will be strengthened. We train the SNNc using the STDP learning rule (Song, Miller, & Abbott, 2000): if neuron j fires before neuron i, then the connection weight from neuron j to neuron i will increase, otherwise the connection from neuron i to neuron j will decrease. This ensures that the time signal inherent in the input spiking trains will be captured by the neuron firing state and manifest as asymmetrical connection weights in the SNNc.

In the SNNc, when a neuron fires, it emits a spike and then its potential gradually reduces over time to reach a zero value. Each neuron connecting to this firing neuron will receive a spike. Because of the emitted spike the potential of each neighboring neuron will increase in proportion to its connection weight to the current firing neuron. The potential of each neuron has a small, constant rate of leakage over time until it reduces to 0. After learning, the connection weights in the SNNc encode temporal relationships from the input

spatio-temporal data (Tu et al., 2014).

8.2.4 Training Stage II:

In this stage, supervised classifier training is carried out. The deSNN (Dhoble et al., 2012; N. Kasabov et al., 2013; N. K. Kasabov, 2014) is used here as an output classifier, because deSNN is computationally efficient and emphasizes the importance of the first spike, which has also been observed to play a significant role in biological systems.

Once the NeuCube(ST) (N. K. Kasabov, 2014) is trained, all connection weights in the SNNc and in the output classification layer are established. For a given new sample without any class label information, the trained NeuCube can be used to predict its class label. For the deSNN classifier, there are two algorithms that can be used to determine the class label of the new sample (Dhoble et al., 2012; N. Kasabov et al., 2013; N. K. Kasabov, 2014). The deSNN classifier uses one of two algorithms to determine the class label of the new sample. One is (Dhoble et al., 2012; N. Kasabov et al., 2013; N. K. Kasabov, 2014) and other is deSNNm (Dhoble et al., 2012; N. Kasabov et al., 2013; N. K. Kasabov, 2014) . We have used the deSNNs algorithm in this work.

8.3 Empirical Results of Age group Classification

Experiments were performed on the publicly available FG-NET (FG-NET, 2002) and MORPH Album 2 (Ricanek Jr & Tesafaye, 2006) databases, both of which are used for benchmarking new methods. We have also extended our experiments to new video (Cerniello, 2013) data for analysis. In all experiments, the size of the SNNc is 1000 neurons, a relatively simple $(10 \times 10 \times 10)$ cube. It is trained and tested using a leave one out method. First, the data is converted into discrete spike trains using the AER encoding method to discretize the continuous signal, following the example of the silicon retina (Hechenbichler & Schliep, 2004)

. The deSNN classifier mentioned previously is used here as an output classifier because deSNN is computationally efficient and emphasizes the importance of the first spike which has been observed to be significant in biological vision systems. We conducted experiments to compare traditional modeling methods (KNN, Naive Bayes and MLP) and our proposed method for age group classification.

We designed three experiments for these baseline algorithms. Note that for these baseline algorithms, the time length of training samples and testing samples have to be the same as these methods cannot tolerate different lengths of feature vectors for training and testing. Table 8.1 shows clearly that the classification achieved with NeuCube significantly outperformed the other techniques. The results clearly indicate that NeuCube with its spatio temporal capability can capture the aging effects more effectively than classical classifiers that do not explicity take into account the time dimension in the data.

Since evolving system data machines (eSTDM) model relationships both between and within spatio-temporal data, even a small amount of input data will be able to trigger the spiking activities in SNNc, giving rise to a more accurate pattern (class) recognition rate from image data.

8.3.1 Age Group Classification with FG-NET

For this experiment we distributed the FG-NET Database into three age groups, the first being the 0-3 age group, the second grouping as 4-16 and the final one as 17-69. We set the sample size to span 5 divisions of time within each age group. We then chose 30 samples from each group, thus enabling 150 images to be selected for each group. For age group 1 the 150 images that were chosen were assigned as class 1. Similarly, we chose 150 images from age group 2 and assigned them as class 2 and did the same for age group 3. For each of the images we take 7 features. Each sample provides a 5x7 matrix with each having 7 features. In all there are 90 samples. The leave-one-person-out testing strategy is used for

the FG-NET database. For comparative purposes we used the same testing strategy as in (Liu et al., 2015).

The data was then fed to NeuCube and Weka. The results of this experiment are given below.

Table 8.1: Age Group Classification using a NeuCube eSTDM in comparison to traditional classifiers: KNN, MLP, NB.

Measure	NeuCube	MLP	KNN	NaiveBayes
Accuracy(%)	98	80	91.1	66.7

We benchmarked our NeuCube results against classical classifiers such as the k nearest neigbor, Multi Layer Perceptron and Naive Bayes.

Table 8.2: Comparative results of age group classification with state-of-the-art age group classifiers

Models	**Classification(%)**
NeuCube	98%
(Liu et al., 2015)	93.5%
(Sai, Wang, & Teoh, 2015)	90%

Table 8.1 shows that NeuCube showed outperformed all of the general purpose classifiers, with the k nearest neighbor performing best amongst this group.

Table 8.2 shows that the purpose-built age classifiers performed better than their general purpose counterparts. However, a substantial gap still exists between them and NeuCube.

8.3.2 Age Group Classification with MORPH

We randomly selected a subset of about 21000 faces from MORPH Album 2 database that contain a mix of black and white, Female and Male faces. Each sample provided a 5x7 matrix with each having 7 features. Our setting for the experiment is similar to that of (Liu et al., 2015) and (G. Guo & Mu, 2011). We divide data into three subsets, S1, S2 and S3. We used S1 for training and tested it with W/S1. We randomly selected a subset W of 21000 images from MORPH database. Then we used S2 for training and tested with W/S2. Finally, we averaged the two results to obtain an overall result.

Table 8.3: Results of Age Group Classification using a NeuCube eSTDM in comparison to traditional classifiers: SVM, MLP, NB.

Measure	NeuCube	MLP	KNN	NaiveBayes
Accuracy(%)	95.00	82.30	89.03	67.60

Table 8.4: Comparative results of age group classification with MORPH Album 2

Models	**Classification(%)**
NeuCube	95%
(Liu et al., 2015)	92.6%
(Sai et al., 2015)	71.2%

Tables 8.3 and 8.4 mirror the results for the FG-NET database, once again underlining the superior performance of NeuCube.

Discussion In Figures 8.2-8.4, blue lines represent positive connections (i.e. those with positive weights) while red lines represent negative connections (those with negative weights). The colour of a neuron signifies its connectivity; the brighter the color, the larger is the size

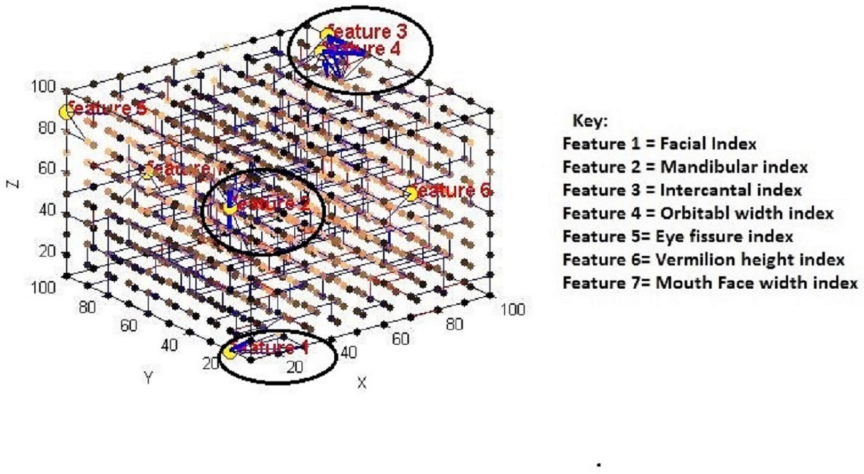


Figure 8.2: Age group (0-3). X,Y and Z are directions

of its neighborhood. The thickness of a line denotes the intensity of a connection between a pair of given neurons. Thus a thick blue line indicates a strong positive connection, while a thin blue line also represents a positive connection but one with a lesser strength.

Figures 8.2-8.4 show that different age group activate different parts of the cube, as indicated by their connectivity.

Figures 8.2 shows that in age group (0-3), the Facial, Intercantal and Orbital width indexes have the strongest level of expression when compared to the other two age groups as they have stronger positive connections when compared to their counterparts in Figures 8.3 and 8.4. Changes in variables in a Spiking Neural Network results in increased spiking activity which in turn spawns positive connections. Thus we can conclude that Cranofacial growth in age group (0-3) occurs at a faster rate than with the other two age groups.

Figure 8.3 shows that in age group (4-16) there is strong connection strength in the Eye fissure and Mandibular index regions due to the presence of thick blue lines. Mandibular and Eye fissure indexes are associated with the eye and chin ratios. This suggests that most

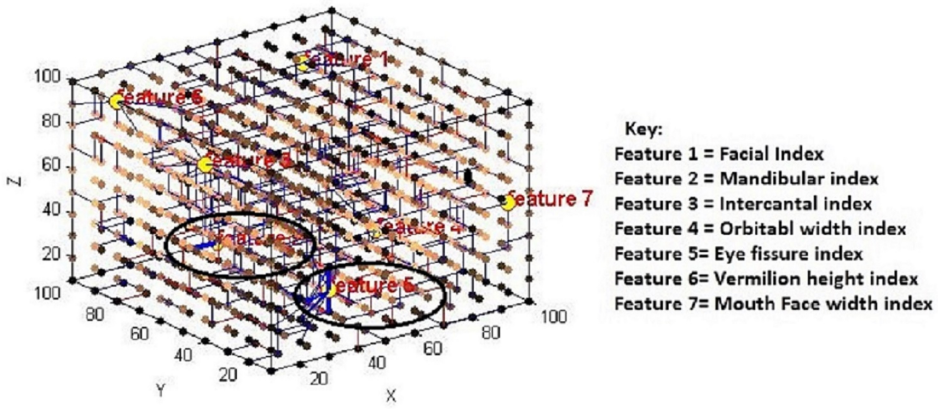


Figure 8.3: Age group (4-16)

changes in the face occur in these features at age group (4-16).

Figure 8.4 shows that in age group (17-69) most of the strong connections seem to occur in the Orbital width, Intercantal, Mandibular and Facial index regions. Even though three of these features, namely the Facial, Intercantal and Orbital indexes, are also expressed in age goup (0-3), the levels of expression is less than that for age group (0-3). Thus we can conclude that while most changes in age group (17-69) occur in these indexes, the rate of change is slower than with age goup (0-3).

Our comparative analysis shows that in all three age groups the Mandibular and Intercantal indexes are active as they are associated with strong spiking activity. The Eye fissure index creates strong spikes for age groups (4-19) and (17-69). The Vermilion height index shows weak connections and a low level of spike activity in age groups (0-3) and (17-69). However, the Vermilion index shows a medium level of spike activity and strong connection weights in age grouping (4-16).

We observe that there is a clear shift away from the Facial, Intercantal, Orbital width and Mouth Face width indexes as markers for the age group (0-3) to Mandibular and Fissure

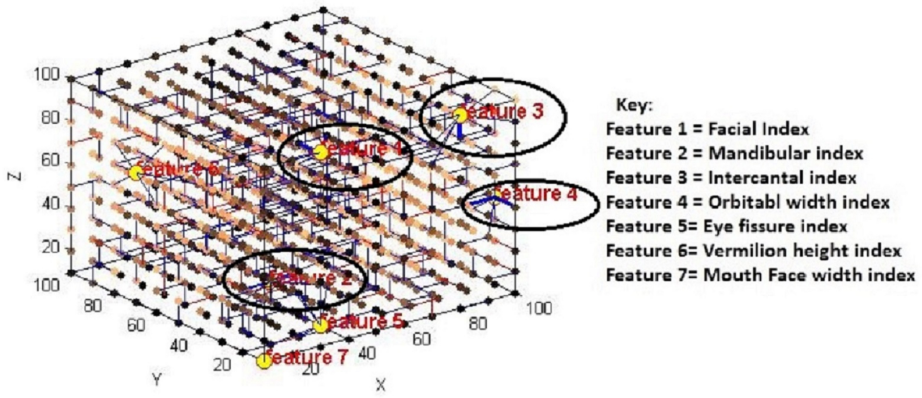


Figure 8.4: Age group (17-69)

indexes for age group (4-16). The fact that different indexes are prominent at different age groups suggests that NeuCube has succeeded in capturing the aging process. This is useful on two different accounts. First, it provides solid evidence that NeuCube is capturing not just the temporal but also spatial signals in the data. Second, from the viewpoint of aging research it provides a useful insight into the changes that take place in the face over time and enables interested researchers into building more useful age estimation models. It also provides the potential to simulate future changes in a person's face over time.

Table 8.5: Feature profile for age group (0-3)

Features	**Influence Level**	**Connections**
Facial	High	Strong
Mandibular	High	Strong
Intercantal	High	Strong
Orbital width	High	Strong
Eye fissure	Low	Weak
Vermilion height	Low	Low
Mouth Face Width	Medium	Strong

Table 8.6: Feature profile for age group (4-16)

Features	**Influence Level**	**Relationship**
Facial	Low	Weak
Mandibular	High	Strong
Intercantal	Medium	Weak
Orbital width	Low	Weak
Eye fissure	High	Strong
Vermilion height	Medium	Weak
Mouth Face Width	Low	Weak

Table 8.7: Feature profile for age group (17-69)

Features	**Influence Level**	**Relationship**
Facial	Medium	Weak
Mandibular	High	Strong
Intercantal	Medium	Strong
Orbital width	High	Strong
Eye fissure	High	Weak
Vermilion height	Low	Weak
Mouth Face Width	Low	Weak

We conclude this section by observing that NeuCube not only outperformed its rivals across both databases in terms of classification accuracy but also yielded useful insights into the aging process.

8.4 Age Group Classification with Video data

A NeuCube aging model is created to classify a given video data into one of three age groups based on its assessed age. The raw data which has been used in this study is from (Cerniello, 2013). It is five minutes of video containing 8943 frames of size (1920×1080) pixels. First the video is converted into greyscale frames. The nose tip of the subject in the image is manually annotated. The purpose was to locate a small region on the face which remains at a fixed distance from the annotated point. That same region is used for all the images in

our study. This region is a part of the texture information of the face image, namely a small part of cheek portion of the face. This is chosen as facial skin that is naturally smooth in youth becomes wrinkled with age, thereby resulting in a change in the textural information present in this area. Based on this assumption 50 pixels are selected from the cheek area of each face image. All frames are divided into three classes. A total of 128 frames are chosen for each sample, for each of the 60 samples. Thus the whole data comprises some 7680 images. The first 20 samples comprise (0-18) youth, the next 20 samples adults (19-35) and the third set of 20 samples represent older (36-70) age persons.

In this experiment the SNNc is trained and tested in a hold out method. Firstly we converted the video data into discrete spike trains using the Temporal Contrast encoding method to discretise the continuous signal, following the example of the silicon retina (Delbruck & Lichtsteiner, 2007). The deSNN classifier mentioned previously is used here as an output classifier due to the fact that deSNN is computationally efficient and emphasises the importance of the first spike, which has been observed in biological vision systems. We conducted experiments to compare between traditional modelling methods (SVM and MLP) and our proposed method for age group classification. We designed two experiments for these baseline algorithms.

Table 8.8: Age group classification accuracy (%) from video data.

Method	**Classification(%)**
SVM	55%
MLP	26%
NeuCube	78%

Note that for these baseline algorithms, the time length of training samples and testing samples have to be the same as these methods cannot tolerate different lengths of feature

vectors for training and testing. It was observed that the classification achieved with NeuCube was better than with other techniques. See Table 8.8 for results. Note that the techniques mentioned (other than NeuCube) do not have the capability of representing the spatio-temporal problem space effectively. These traditional techniques are only suitable for static data within a given time segment. Since an eSTDM models the relationships between and within spatio-temporal data, even a small input data will be able to trigger the spiking activities in SNNc, for an accurate pattern (class) recognition from video data.

8.5 Empirical Results of Gender Recognition

8.5.1 Gender Recognition with FG-NET

For this experiment we use the FG-NET Database. We model Gender recognition as a two-class classification problem. We assign the gender label through a visual inspection of the data, with class 1 as male and class 2 as female. We also used the classical classifiers from Weka as in the first experiment for comparative analysis. For each of the images we take 7 features. Each sample provides a 5x7 matrix, ie 5 images, each having 7 features. In all there are 172 samples. The leave-one-person-out testing strategy is used for FG-NET. Table 8.9 shows the results of the age group classification process.

Table 8.9: Gender Classification on FG-NET

Measure	NeuCube	MLP	KNN	NaiveBayes
Accuracy(%)	95.00	76.10	83.00	62.79

Tables 8.9 and 8.10 show clearly that NeuCube has better classification performance when compared to all of its rivals.

Table 8.10: Age group classification on FG-NET

Models	**Classification(%)**
NeuCube (G. Guo & Mu, 2011)	95% 87.20%

8.5.2 Gender Recognition with FG-NET in a younger age group (0-18)

For this experiment we use the FG-NET Database. For each of the images we take 7 features. Each sample provides a 5x7 matrix ie 5 images, with each having 7 features. In all there are 120 samples. The leave-one-person-out testing strategy was used for FG-NET. Table 8.11 shows results of gender recognition in the younger age group.

Table 8.11: Gender Classification for younger age group (0-18)

Measure	NeuCube	MLP	KNN	NaiveBayes
Accuracy(%)	85.00	72.60	70.80	67.3

8.5.3 Gender Recognition with MORPH

For this experiment we use the MORPH Album 2 Database. We model Gender recognition as a two-class classification problem. We use the same testing strategy as used in (G. Guo & Mu, 2011; Liu et al., 2015). This data is then fed into the NeuCube. We use the same classical classifiers from Weka as we did in the first experiment for comparative analysis. For each of the images we take 7 features. Each sample provides a 5x7 matrix ie 5 images, with each having 7 features. In all there are 3378 samples.

Table 8.12: Results of Gender Classification using a NeuCube eSTDM in comparison with traditional techniques: KNN, MLP, NB

Measure	NeuCube	MLP	KNN	NaiveBayes
Accuracy(%)	99.00	76.30	88.00	66.00

Table 8.13: Comparative results of age group classification with FG-NET database

Models	**Classification(%)**
NeuCube (G. Guo & Mu, 2011)	99% 98.20%

Tables 8.12 and 8.13 show clearly that NeuCube has better classification performance when compared to all of its rivals.

8.5.4 Gender Recognition with MORPH in a younger age group (16-20)

In all there are 420 samples that we use for this experiment. Each sample provides a 5x7 matrix with each image having 7 features. We use the same testing strategy as used in (G. Guo & Mu, 2011; Liu et al., 2015). Table 8.14 shows results of gender recognition in the younger age groups.

It was observed that the gender recognition classification achieved with NeuCube was better than with other traditional techniques for the younger age groups.

Discussion In Figures 8.5 and 8.6, line colour, thickness and coloring of neurons have similar meaning to that described above.

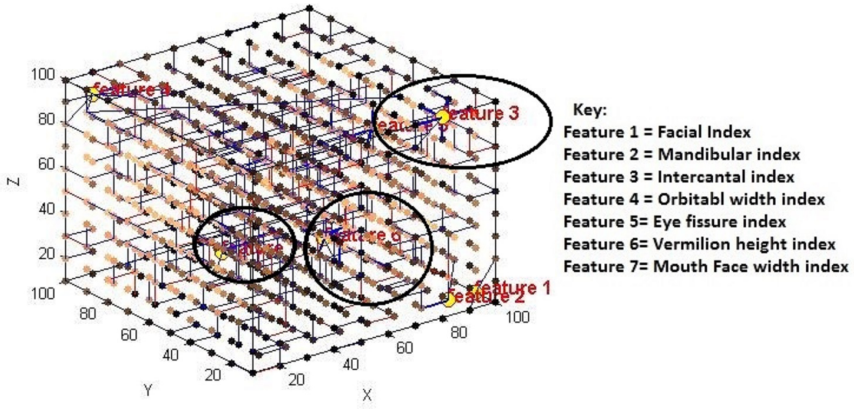


Figure 8.5: Male

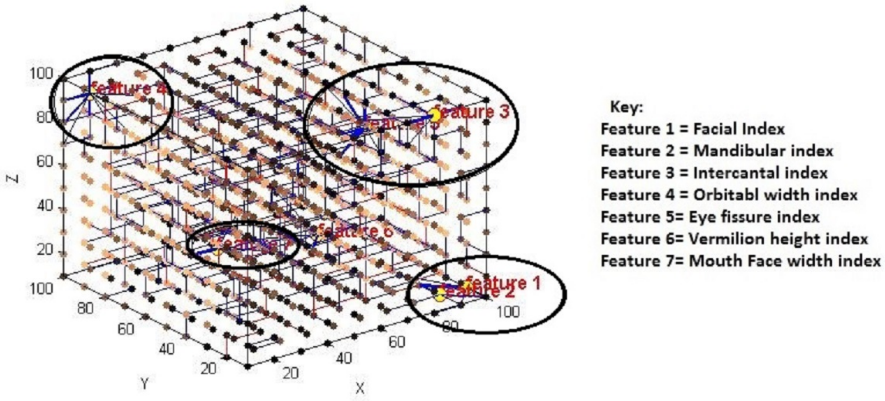


Figure 8.6: Female

Table 8.14: Results of Gender Classification at a younger age group (16-20) using a NeuCube eSTDM in comparison with traditional techniques: KNN, MLP, NB

Measure	NeuCube	MLP	KNN	NaiveBayes
Accuracy(%)	87.0	68.3	79.3	63.8

By visualization of the cube in Figure 8.5 we can see that the Intercantal, Vermilion height, Eye fissure and Mouth Face Width indexes have strong positive connections. However, the Mandibular, Facial and orbital indexes show weak connections caused by less spiking activity.

In Figure 8.6 we see that all features generate high spikes and stronger connections except for the Vermilion height index. This is typical for the female specimens in the data sample.

By conducting a comparative analysis we observe that for both genders the Intercantal, Eye Fissure and Mouth Face Width indexes show strong, positive connections, implying that more changes would occur in eyes, mouth and bone structure between eyes.

The Facial, Mandibular, Orbital width and Vermilion height indexes behave differently across the genders, implying that a gender aging pattern exists. This further validates the point that gender also effects the face recognition process. Therefore, it is recommended that gender-specific aging models be built for age invariant face recognition.

By visualization we can also see another interesting effect, facial index, mandibular index, intercantal index and orbital index show strong connections for female gender as compared to male gender. This suggests that female gender ages more quickly as compared to male gender.

In these experiments we can see that the NeuCube has the capability of providing better accuracy as compared to the traditional methods. Its main strength lies in three dimensional

connections and their weights and the ability of learning spatio temporal aging process.

Table 8.15: Feature profile (Male)

Features	**Influence Level**	**Connections**
Facial	Low	Weak
Mandibular	Low	Weak
Intercantal	High	Strong
Orbital width	Low	Weak
Eye fissure	High	Strong
Vermilion height	High	Strong
Mouth Face Width	High	Strong

Table 8.16: Feature profile (Female)

Features	**Influence Level**	**Connections**
Facial	High	Strong
Mandibular	High	Strong
Intercantal	High	Strong
Orbital width	High	Strong
Eye fissure	High	Strong
Vermilion height	Low	Low
Mouth Face Width	High	strong

8.6 Conclusion

In this chapter we have shown two inter-related problems in face recognition using the NeuCube computational platform. The FG-NET, MORPH and video(Cerniello, 2013) databases were used and Anthropometric features were extracted from landmark points on the face. These features enabled learning of spatio temporal relationships. The Weka machine learning workbench was used to compare the performance of traditional classifiers such as the K nearest neighbor (KNN), Multi-LayerPerceptron (MLP) and Naive Bayes with NeuCube. This research has revealed that NeuCube has the capability of giving better performance for gender recognition and age group classification because it has the capacity to learn both spatial and temporal relationships simultaneously.

Our empirical results show that NeuCube performed consistently better across both

problem types that we investigated. We explored the temporal relationships between different Anthropometric features. The aging process causes significant alterations on the human face, thus affecting the long term performance of face authentication systems. Our research also revealed that different Fiducial markers are prominent at different age groups, thus offering the possibility of estimating aging effects for future points in a given person's timeline. However, for age simulation to be successful such markers may need to be combined with other person-specific information such as lifestyle and environment variables that could also impact on shaping a person's facial profile over time.

Chapter 9

Conclusion and Further Research

9.1 Conclusion

This research has presented a novel method for age invariant face recognition and inter-related problems to face recognition viz age group classification and gender recognition. The developed models could also be used for knowledge extraction and information extraction on the aging process. To the best of my knowledge, this study is the first comprehensive study of using spatio temporal modeling methods for modeling aging in biometrics. During the course of this study different aging models have been developed at different levels of data granularity. These models have been used for knowledge extraction and for age invariant face recognition. The NeuCube aging model based on the spiking neural network architecture has been used for gender recognition and age group classification.

On the basis of research objectives we outlined our research questions in Chapter 1. These research questions have been investigated by utilising the two publically available databases, viz MORPH and FG-NET. The questions that we investigated during this research are:

1. The development of comprehensive methods for age invariant face recognition that result in better accuracy and efficiency than current state-of-the-art methods have been discussed in Chapters 3, 4 and 5. Chapter 3 presents global and personalized

models for capturing aging patterns with texture and shape information. A Personalized model covers the individual aging patterns while a Global model captures general aging patterns across the entire population captured in a database. A novel de-aging factor has been introduced that de-ages each individual's age in a given probe set. The personalized model employes KNN and regression for constructing the aging models. It also uses a Bayesian method for autmomatic selection of an aging subspace. The developed models are then put in place for age invariant face recognition. Chapter 4 presents a local model for texture and shape information which captures homogenity within a sub-population. It employs dynamic time warping and the k means clustering algorithm for constucting clusters that repsrent the sub-populations. Chapter 5 presents the application of an Integrated Multi-Model Framework to the age invariant recognition problem. The integrated framework utilizes the proposed global (Chapter 3), personalized (Chapter 3), and local (Chapter 4) aging models for texture and shape information and exploits synergy that exists between the three types of models for age invariant face recognition. Our results show that our proposed models performed better as compared to state of art approaches.

2. The development of a framework that incorporates both texture and shape features at system level while effectively dealing with conflicts between the two model types was presented in Chapter 6. It discusses a composite model for resolving conflictx between texture and shape models built at different levels of data granularity. It uses a decision tree method for resolving the conflicts between different types of aging models. It also discusses about different models and explains which model is useful at different situations.

3. A craniofacial growth model has been developed that maximizes predictive power, as measured by recognition accuracy. This craniofacial growth model can also be used

for face simulation by incorporating texture information.

4. Chapter 7 presented the craniofacial growth model. Aging effects have been captured at different time points by the proposed aging models. The Chapter demonstrated that the two genders age differently. The shape of a person face changes in the horizontal and vertical dimensions but that rate of change is personalized to that person. Moreoever, at certain time points different types of changes were observed and analyzed.

5. The development of efficient methods for inter-related problems with face recognition, viz age group clasification and gender recognition, in order to increase the accuracy have been discussed in Chapter 8.

   The results suggested that NeuCube's aging models performed better as compared to state of art algorithms.

6. Knowledge extraction was carried out with the proposed methods for gender recognition and age group classification.

   The NeuCube aging models were tested on both publically available datasets. The knowledge discovered shows that anthropometric features have temporal relationships with each other. The analsyis showed that, at different age groups, feature profile behaviour specific to that age grouping was successfully captured by NeuCube.

The next section will discuss about our accomplishments during this research.

9.2 Accomplishments

In this research we presented shape and texture models at different levels of data granularity. The Personalized model, where a model is built for each individual at the lowest level of

granularity yielded the best Rank 1 recognition rate. This indicates that aging to a large extent is an individualized process and this finding is in agreement with other studies.

Our method of dividing the whole database into subspaces and using the Naive Bayesian classifier to narrow down the search space also proved to be a success. This limits the search space and improves recognition accuracy.

In the case of the Local model we further reduced the search space by first dividing the whole database into sub-populations (clusters) and then considered subspaces within each sub-population. The Local model traded off accuracy with recognition time. Reduction in recognition time follows directly from restricting the search space to a cluster rather than the entire dataset. Experimentation with the databases resulted in a reduction of recognition time by a factor of 60% in relation to the Personalized model approach but a reduction in Rank 1 accuracy also resulted due to the fact that image recognition decisions are made on localized data rather than on a global dataset.

The Integrated model which consisted of a combination across granularity level (Global, Local, Personalized) resulted in the highest accuracy because of the synergy that existed between the models. The increased accuracy did come at the cost of extra computation time, increasing the recognition time 35% in relation to the Personalized model approach.

Another idea that we explored in this research was model composition between models built from shape and texture features. It turned out that model composition, similar to model integration, resulted in better accuracy. The final conclusion of the study was that the optimal modeling approach was to compose shape models built by integrating across data granularity levels with their corresponding texture counterparts using the same data granularity scheme.

A brief comparison also suggested which of the models can be used depending on the end user's requirements. For highest accuracy the Integrated Model is suggested and when time is at a premium, the Local Model is a good compromise as it has a lower recognition

time than the Personalized model while having better accuracy than the Global model.

Craniofacial growth model shows that our aging models are also capable of forward traversal through time. Forward traversal bring with it two attractive benefits. The first is automatic update of the image database. Over a period of time an image gallery containing older images of a person becomes less useful for recognition purposes as research has shown that longer the time gap between the probe image and the target image, the lower is the Rank 1 accuracy in general (Park et al., 2010). The solution to this problem either involves obtaining newer images from the subjects concerned which may not be practicable as some subjects may not be available. A more practical alternative would be an automatic refresh based on a computed trajectory provided by the aging functions proposed in this research. The second benefit lies in age simulation whereby the appearance of given person at a future point in time is required for applications as face animation or for making decisions at a future point in time on plastic surgery alternatives.

We also researched two inter-related problems in face recognition using the NeuCube computational platform. The well-known FG-NET and MORPH image galleries were used and Anthropometric features were extracted from landmark points on the face. These features enabled learning of spatio temporal relationships. The Weka machine learning workbench was used to compare the performance of traditional classifiers such as the K nearest neighbor (Knn), Multi-LayerPerceptron (MLP) and Naive Bayes with NeuCube. This research has revealed that NeuCube has the capability of giving better performance for gender recognition and age group classification because it has the capacity to learn both spatial and temporal relationships simultaneously.

Our empirical results show that NeuCube performed consistently better across both problem types that we investigated. We explored the temporal relationships between different Anthropometric features. The aging process causes significant alterations to the human face, thus affecting the long-term performance of face authentication systems. Our research

also revealed that different Fiducial markers are prominent at different age groups, thus offering the possibility of estimating aging effects for future points in a given person's timeline. However, for age simulation to be successful such markers may need to be combined with other person-specific information such as lifestyle and environment variables that could also impact on shaping a person's facial profile over time.

In conclusion we observe that NeuCube provides a good platform for aging research, whether the objective is age estimation or estimating change in facial features over time, thus opening up future research opportunities in these two key areas of aging research.

9.3 Limitations of Research

9.3.1 Lack of Database

Since aging is a spatio temporal process it is quite difficult to collecta comprehensive training dataset because of temporal nature of the domain in our problem space. It affects the quality of aging models that are generated because they are highly dependent on volume and quality of data available for training purposes. The ideal database should have sufficient information regarding lifestyle, economic conditions, bio markers such as genes and blood group, as well as social data such as ethnicity. Moreoever, databases should have minimum number of aberrations in images with respect to confounding factors such as allumination, pose, expression, occlusion and facial expressions. Such datasets will help in better understanding of the aging process.

9.3.2 Dependency on human aging expert

The generative methods are highly dependent on age of the person. Currently, it is assumed that a human aging expert helps to identify the age of a probe image. It is more realistic to use an age estimation method and incorporate it into the system. Moreover, effects of texture and shape on age estimation would also make for an interesting study.

There are further limitations which will be addressed in the future work section.

9.4 Directions of Future Research

9.4.1 Dynamic deviation factor

We can incorporate a dynamic deviation factor that changes with age. One direction is to use the Markovian property(Rao & Kshirsagar, 1978; Shyama & Linda, 2016) and incorporate hidden markov process techniques(Uiboupin, Rasti, Anbarjafari, & Demirel, 2016; B. Wu, Hu, & Ji, 2016).

Each person can be represented by hidden markov model $\lambda(A, B, \pi)$ where A represents the state transition probability matrix , B represents observation probability matrix and π initial state distribution matrix. After the initialization the parameters would be updated using the Baum-Welch algorithm(Welch, 2003) and optimal path can be found by using Viterbi algorithm(Forney, 1973). This optimal path information can be used in place of the deviation factor formulation currently formulated in Chapter 3.

At each ageband three possible states of a person can occur viz growth, persistence or decay. Therefore state refers to change status within the ageband (growth, persistent or decay) as manifested by fiducial feature value. The transition probabilities can be learned by using training data. A given person's aging trajectory will be determined by the status taken within each age band. Since it is possible for a person to transit from one state to another within an age band, multiple aging paths could be constructed for a given person. The Viterbi algorithm could then be used to construct an aging trajectory that corresponds to the path with the maximum liklihood given the set of all poissble paths for an individual.

Also different machine learning algorithms such as genetic algorithm (Janikow, 1993) can be tested for better representation of deviation factor.

It has been claimed that via the operations of selection, crossover, and mutation the Genetic Algorithm (GA) will converge over successive generations towards the global (or

near global) optimum. The fitness function can be determined by using the Euclidean distance between the probe image feature vector and the nearest matching training set feature vector.

The basic objective is to minimize the error in fitness objective function. Therefore, the solution returned by the GA can be used as an alternative formulation of the deviation factor vector.

9.4.2 Multivariate tecnique for age invariant face recognition

It would be interesting to use multivariate regression(Scott, 2015; Harrell, 2015; Chatterjee & Hadi, 2015; Wegbreit et al., 2015; Dhall, Goecke, Joshi, Hoey, & Gedeon, 2016) and incorporate other attributes related to aging such as ethnicity, gender, blood group, life style information etc. Analysis of these attributes would give important insights into the aging process.

9.4.3 Aging models for face simulation

The proposed models can be used for building a synthesized face(Tang & Wang, 2003; Wang & Tang, 2009; Aizawa, Harashima, & Saito, 1989). We would need to add texture information to craniofacial shape growth model as well as wrinkles and scars on the face using warping techniques(Wolberg, 1990; Brown, 1992; Glasbey & Mardia, 1998; Pishchulin, Gass, Dreuw, & Ney, 2012).

9.4.4 NeuCube aging models for age estimation

One direction that can be explored is to construct NeuCube aging models for each feature separately and then fuse the different models to build a composite model(Alvi & Pears, 2016) of the simulated face. Modelling each feature separately is attractive as it allows each feature to express its individual contribution to the aging process to the maximum possible

extent without interference from the other features(Alvi & Pears, 2015). Our future research will explore this approach in depth.

9.4.5 NeuCube aging models for real time use

The NeuCube aging models implementation can be done for real time processing face recognition. As a NeuCube simulator is available in PyNN, along with Java and Matlab(N. Kasabov et al., 2016) this makes it possible for a direct implementation of such models on all available neuromorphic hardware platforms for very fast processing on large volumes of data in an online, real time mode. Computational platforms such as the Manchester SpiNNaker(Furber, Galluppi, Temple, & Plana, 2014; Mendat, Chin, Furber, & Andreou, 2016) and the IBM TrueNorth(Merolla et al., 2014) along with hybrid neuromorphic chips(Indiveri et al., 2011), of thousands and millions of spiking neurons with very low energy consumption, can now be used for real time age group classification and age estimation from large databases(Bose, Kasabov, Bruzzone, & Hartono, 2016). In this respect the proposed method is the first to enable a direct use of neuromorphic hardware for such data modeling and analysis.

9.4.6 3D models for aging

Moreover, 3D anthropometric(Luximon, Ball, & Justice, 2012; Seo, Song, Kim, & Kim, 2016; Horprasert, Yacoob, & Davis, 1997; Evison & Bruegge, 2016) models can be created for better representation of shape. More sophisticated texture features such as Local Binary pattern(Ojala, Pietikainen, & Maenpaa, 2002; Mirmehdi, Xie, & Suri, 2008; Z. Guo, Zhang, & Zhang, 2010) and scale invariant feature transforms(Lowe, 1999; Juan & Gwun, 2009; Lowe, 2004) can be extracted from the face image. It would be an interesting study to fuse generative and non generative approaches for aging problems. It is expected that they would also be helpful for understanding of aging process.

This is due to the fact that fusion could help to study the effects of features that are

variant with age as well as those features that remain relatively invariant during the lifespan of a person.

References

Abate, A. F., Nappi, M., Riccio, D., & Sabatino, G. (2007). 2d and 3d face recognition: A survey. *Pattern Recognition Letters*, *28*(14), 1885–1906.

Abdullah, M. F. A., Sayeed, M. S., Muthu, K. S., Bashier, H. K., Azman, A., & Ibrahim, S. Z. (2014). Face recognition with symmetric local graph structure (slgs). *Expert Systems with Applications*, *41*(14), 6131–6137.

Ahonen, T., Hadid, A., & Pietikainen, M. (2006). Face description with local binary patterns: Application to face recognition. *IEEE transactions on pattern analysis and machine intelligence*, *28*(12), 2037–2041.

Aizawa, K., Harashima, H., & Saito, T. (1989). Model-based analysis synthesis image coding (mbasic) system for a person's face. *Signal Processing: Image Communication*, *1*(2), 139–152.

Alvi, F. B., & Pears, R. (2015, Nov). Use of spatio-temporal modeling for age invariant face recognition. In *2015 international conference on image and vision computing new zealand (ivcnz)* (p. 1-6). doi: 10.1109/IVCNZ.2015.7761566

Alvi, F. B., & Pears, R. (2016). A composite spatio-temporal modeling approach for age invariant face recognition. *Expert Systems with Applications*.

Amor, N. B., Benferhat, S., & Elouedi, Z. (2004). Naive bayes vs decision trees in intrusion detection systems. In *Proceedings of the 2004 acm symposium on applied computing* (pp. 420–424).

Baruch, I., & Stoyanov, I. (1995). Global model of neural networks, stability and learning. In *Electrical and electronics engineers in israel, 1995., eighteenth convention of* (pp. 4–3).

Belhumeur, P. N., Hespanha, J. P., & Kriegman, D. J. (1997). Eigenfaces vs. fisherfaces: Recognition using class specific linear projection. *IEEE Transactions on pattern analysis and machine intelligence*, *19*(7), 711–720.

Biswas, S., Aggarwal, G., Ramanathan, N., & Chellappa, R. (2008). A non-generative approach for face recognition across aging. In *Biometrics: Theory, applications and systems, 2008. btas 2008. 2nd ieee international conference on* (pp. 1–6).

Bolme, D. S. (2003). *Elastic bunch graph matching* (Unpublished doctoral dissertation). Colorado State University.

Bose, P., Kasabov, N. K., Bruzzone, L., & Hartono, R. N. (2016). Spiking neural networks for crop yield estimation based on spatiotemporal analysis of image time series. *IEEE Transactions on Geoscience and Remote Sensing*, *54*(11), 6563–6573.

Bowyer, K. W., Chang, K., & Flynn, P. (2006). A survey of approaches and challenges in 3d and multi-modal 3d+ 2d face recognition. *Computer vision and image understanding*, *101*(1), 1–15.

Brown, L. G. (1992). A survey of image registration techniques. *ACM computing*

surveys (CSUR), *24*(4), 325–376.

Bruce, V., & Young, A. (1986). Understanding face recognition. *British journal of psychology*, *77*(3), 305–327.

Canny, J. (1986). A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*(6), 679–698.

Champod, C. (2015). Forensic applications, overview. In S. Z. Li & A. K. Jain (Eds.), *Encyclopedia of biometrics* (pp. 705–712). Boston, MA: Springer US. Retrieved from `http://dx.doi.org/10.1007/978-1-4899-7488-4_100` doi: 10.1007/978-1-4899-7488-4_100

Chatterjee, S., & Hadi, A. S. (2015). *Regression analysis by example*. John Wiley & Sons.

Cootes, T. F., Edwards, G. J., Taylor, C. J., et al. (1999). Comparing active shape models with active appearance models. In *Bmvc* (Vol. 99, pp. 173–182).

Delac, K., Grgic, M., & Grgic, S. (2005a). Generalization abilities of appearance-based subspace face recognition algorithms. In *Intl. workshop on systems, signals and image processing, chalkida, greece* (pp. 273–276).

Delac, K., Grgic, M., & Grgic, S. (2005b). Independent comparative study of pca, ica, and lda on the feret data set. *International Journal of Imaging Systems and Technology*, *15*(5), 252–260.

Déniz, O., Castrillon, M., & Hernández, M. (2003). Face recognition using independent component analysis and support vector machines. *Pattern recognition letters*, *24*(13), 2153–2157.

Dhall, A., Goecke, R., Joshi, J., Hoey, J., & Gedeon, T. (2016). Emotiw 2016: Video and group-level emotion recognition challenges. In *Proceedings of the 18th acm international conference on multimodal interaction* (pp. 427–432).

Dhoble, K., Nuntalid, N., Indiveri, G., & Kasabov, N. (2012). Online spatio-temporal pattern recognition with evolving spiking neural networks utilising address event representation, rank order, and temporal spike learning. In *Neural networks (ijcnn), the 2012 international joint conference on* (pp. 1–7).

Ding, C., Choi, J., Tao, D., & Davis, L. S. (2016). Multi-directional multi-level dual-cross patterns for robust face recognition. *IEEE transactions on pattern analysis and machine intelligence*, *38*(3), 518–531.

Ding, C., & Tao, D. (2016). A comprehensive survey on pose-invariant face recognition. *ACM Transactions on Intelligent Systems and Technology (TIST)*, *7*(3), 37.

Dunstone, T., & Yager, N. (2008). *Biometric system and data analysis: Design, evaluation, and data mining*. Springer Science & Business Media.

Evison, M. P., & Bruegge, R. W. V. (2016). *Computer-aided forensic facial comparison*. CRC Press.

Farkas, L. G. (1994). *Anthropometry of the head and face*. Raven Pr.

Farkas, L. G., & Munro, I. R. (1987). *Anthropometric facial proportions in medicine.* Charles C. Thomas Publisher.

FG-NET. (2002). *Fg-net database.* Retrieved from `http://www-prima.inrialpes.fr/FGnet/`

Forney, G. D. (1973). The viterbi algorithm. *Proceedings of the IEEE*, *61*(3), 268–278.

Friedl, M. A., & Brodley, C. E. (1997). Decision tree classification of land cover from remotely sensed data. *Remote sensing of environment*, *61*(3), 399–409.

Fu, Y., & Huang, T. S. (2008). Human age estimation with regression on discriminative aging manifold. *Multimedia, IEEE Transactions on*, *10*(4), 578–584.

Furber, S. B., Galluppi, F., Temple, S., & Plana, L. A. (2014). The spinnaker project. *Proceedings of the IEEE*, *102*(5), 652–665.

Furey, T. S., Cristianini, N., Duffy, N., Bednarski, D. W., Schummer, M., & Haussler, D. (2000). Support vector machine classification and validation of cancer tissue samples using microarray expression data. *Bioinformatics*, *16*(10), 906–914.

Gates, K. (2011). *Our biometric future: Facial recognition technology and the culture of surveillance.* NYU Press.

Geng, X., Zhou, Z.-H., & Smith-Miles, K. (2007). Automatic age estimation based on facial aging patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *29*(12), 2234–2240.

Gheissari, N., Sebastian, T. B., & Hartley, R. (2006). Person reidentification using spatiotemporal appearance. In *2006 ieee computer society conference on computer vision and pattern recognition (cvpr'06)* (Vol. 2, pp. 1528–1535).

Glasbey, C. A., & Mardia, K. V. (1998). A review of image-warping methods. *Journal of applied statistics*, *25*(2), 155–171.

Gong, D., Li, Z., Lin, D., Liu, J., & Tang, X. (2013). Hidden factor analysis for age invariant face recognition. In *Proceedings of the ieee international conference on computer vision* (pp. 2872–2879).

Gong, D., Li, Z., Tao, D., Liu, J., & Li, X. (2015). A maximum entropy feature descriptor for age invariant face recognition. In *Proceedings of the ieee conference on computer vision and pattern recognition* (pp. 5289–5297).

Goshtasby, A. A. (2012). *Image registration: Principles, tools and methods.* Springer Science & Business Media.

Guo, G., Fu, Y., Dyer, C. R., & Huang, T. S. (2008). Image-based human age estimation by manifold learning and locally adjusted robust regression. *Image Processing, IEEE Transactions on*, *17*(7), 1178–1188.

Guo, G., & Mu, G. (2011). Simultaneous dimensionality reduction and human age estimation via kernel partial least squares regression. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 657–664. doi: 10.1109/CVPR.2011.5995404

Guo, Z., Zhang, L., & Zhang, D. (2010). A completed modeling of local binary pattern operator for texture classification. *IEEE Transactions on Image Processing*, *19*(6), 1657–1663.

Harrell, F. (2015). *Regression modeling strategies: with applications to linear models, logistic and ordinal regression, and survival analysis*. Springer.

Haykin, S. S., Haykin, S. S., Haykin, S. S., & Haykin, S. S. (2009). *Neural networks and learning machines* (Vol. 3). Pearson Upper Saddle River, NJ, USA:.

Hechenbichler, K., & Schliep, K. (2004). *Weighted k-nearest-neighbor techniques and ordinal classification* (Tech. Rep.). Discussion paper//Sonderforschungsbereich 386 der Ludwig-Maximilians-Universität München.

Heitmeyer, R. (2000). Biometric identification promises fast and secure processing of airline passengers. *ICAO Journal*, *55*(9), 10–11.

Horprasert, T., Yacoob, Y., & Davis, L. S. (1997). Computing 3d head orientation from a monocular image sequence. In *25th annual aipr workshop on emerging applications of computer vision* (pp. 244–252).

Hwang, Y.-C. (2009). *Local and personalised models for prediction, classification and knowledge discovery on real world data modelling problems* (Unpublished doctoral dissertation). Auckland University of Technology.

Indiveri, G., Linares-Barranco, B., Hamilton, T. J., Van Schaik, A., Etienne-Cummings, R., Delbruck, T., ... others (2011). Neuromorphic silicon neuron circuits. *Frontiers in neuroscience*, *5*, 73.

İşcan, M. Y., & Helmer, R. P. (1993). *Forensic analysis of the skull: craniofacial analysis, reconstruction, and identification*. John Wiley & Sons Inc.

Jain, A., Flynn, P., & Ross, A. A. (2007). *Handbook of biometrics*. Springer Science & Business Media.

Jain, A. K., & Li, S. Z. (2011). *Handbook of face recognition*. Springer.

Jain, A. K., Nandakumar, K., & Ross, A. (2016). 50 years of biometric research: Accomplishments, challenges, and opportunities. *Pattern Recognition Letters*.

Jana, R., Datta, D., & Saha, R. (2013). Age group estimation using face features. *International Journal of Engineering and Innovative Technology*, *3*(2).

Janikow, C. Z. (1993). A knowledge-intensive genetic algorithm for supervised learning. *Machine learning*, *13*(2-3), 189–228.

Juan, L., & Gwun, O. (2009). A comparison of sift, pca-sift and surf. *International Journal of Image Processing (IJIP)*, *3*(4), 143–152.

Kanade, T. (1973). Picture processing system by computer complex and recognition of human faces. *Doctoral dissertation, Kyoto University*, *3952*, 83–97.

Kasabov, N. (2001). Evolving fuzzy neural networks for supervised/unsupervised online knowledge-based learning. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, *31*(6), 902–918.

Kasabov, N. (2007a). *Evolving connectionist systems: the knowledge engineering*

approach. Springer Science & Business Media.

Kasabov, N. (2007b). Global, local and personalised modeling and pattern discovery in bioinformatics: An integrated approach. *Pattern Recognition Letters*, *28*(6), 673–685.

Kasabov, N. (2012). Neucube evospike architecture for spatio-temporal modelling and pattern recognition of brain signals. In *Artificial neural networks in pattern recognition* (pp. 225–243). Springer.

Kasabov, N., Dhoble, K., Nuntalid, N., & Indiveri, G. (2013). Dynamic evolving spiking neural networks for on-line spatio-and spectro-temporal pattern recognition. *Neural Networks*, *41*, 188–201.

Kasabov, N., Feigin, V., Hou, Z.-G., Chen, Y., Liang, L., Krishnamurthi, R., ... Parmar, P. (2014). Evolving spiking neural networks for personalised modelling, classification and prediction of spatio-temporal patterns with a case study on stroke. *Neurocomputing*, *134*, 269–279.

Kasabov, N., Scott, N. M., Tu, E., Marks, S., Sengupta, N., Capecci, E., ... others (2015). Evolving spatio-temporal data machines based on the neucube neuromorphic framework: Design methodology and selected applications. *Neural Networks*.

Kasabov, N., Scott, N. M., Tu, E., Marks, S., Sengupta, N., Capecci, E., ... others (2016). Evolving spatio-temporal data machines based on the neucube neuromorphic framework: design methodology and selected applications. *Neural Networks*, *78*, 1–14.

Kasabov, N. K. (2014). Neucube: A spiking neural network architecture for mapping, learning and understanding of spatio-temporal brain data. *Neural Networks*, *52*, 62–76.

Kass, M., Witkin, A., & Terzopoulos, D. (1988). Snakes: Active contour models. *International journal of computer vision*, *1*(4), 321–331.

K.jain, A. (2009). *Biometric recognition: How do i know who you are?* Retrieved from `http://biometrics.cse.msu.edu/Presentations/JainBiometricsTalk12May09.pdf`

Klare, B., & Jain, A. K. (2011). Face recognition across time lapse: On learning feature subspaces. In *International joint conference on biometrics (ijcb)* (pp. 1–8).

Kuncheva, L. I. (2004). *Combining pattern classifiers: methods and algorithms*. John Wiley & Sons.

Kwon, Y. H., & da Vitoria Lobo, N. (1999). Age classification from facial images. *Computer Vision and Image Understanding*, *74*(1), 1–21.

Lanitis, A., Taylor, C. J., & Cootes, T. F. (1999). Modeling the process of ageing in face images. In *The proceedings of the seventh ieee international conference on computer vision* (Vol. 1, pp. 131–136).

Li, Z., Gong, D., Li, X., & Tao, D. (2016, May). Aging face recognition: A hierarchical learning model based on local patterns selection. *IEEE Transactions on Image Processing*, *25*(5), 2146-2154. doi: 10.1109/TIP.2016.2535284

Li, Z., Park, U., & Jain, A. K. (2011). A discriminative model for age invariant face recognition. *Information Forensics and Security, IEEE Transactions on*, *6*(3), 1028–1037.

Liaw, A., & Wiener, M. (2002). Classification and regression by randomforest. *R news*, *2*(3), 18–22.

Ling, H., Soatto, S., Ramanathan, N., & Jacobs, D. W. (2007). A study of face recognition as people age. In *Ieee 11th international conference on computer vision (iccv)* (pp. 1–8).

Liu, K.-H., Yan, S., & Kuo, C.-C. J. (2015). Age Estimation via Grouping and Decision Fusion. *IEEE Transactions on Information Forensics and Security*, *10*(11), 2408–2423. Retrieved from `http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=7173035` doi: 10.1109/TIFS.2015.2462732

Lowe, D. G. (1999). Object recognition from local scale-invariant features. In *Computer vision, 1999. the proceedings of the seventh ieee international conference on* (Vol. 2, pp. 1150–1157).

Lowe, D. G. (2004). Distinctive image features from scale-invariant keypoints. *International journal of computer vision*, *60*(2), 91–110.

Lucks, M., Oki, N., & RamirezAngulo, J. (1999). A radial basis function network (rbfn) for function approximation. In *42nd midwest symposium on circuits and systems, proceedings, vols 1 and 2* (pp. 1099–1101).

Luximon, Y., Ball, R., & Justice, L. (2012). The 3d chinese head and face modeling. *Computer-Aided Design*, *44*(1), 40–47.

Marin, F., Garcia-Lagos, F., Joya, G., & Sandoval, F. (2002). Global model for short-term load forecasting using artificial neural networks. *IEE Proceedings-Generation, Transmission and Distribution*, *149*(2), 121–125.

Martínez, A. M. (2002). Recognizing imprecisely localized, partially occluded, and expression variant faces from a single sample per class. *IEEE Transactions on Pattern analysis and machine intelligence*, *24*(6), 748–763.

Mendat, D. R., Chin, S., Furber, S., & Andreou, A. G. (2016). Neuromorphic sampling on the spinnaker and parallella chip multiprocessors. In *2016 ieee 7th latin american symposium on circuits & systems (lascas)* (pp. 399–402).

Merolla, P. A., Arthur, J. V., Alvarez-Icaza, R., Cassidy, A. S., Sawada, J., Akopyan, F., ... others (2014). A million spiking-neuron integrated circuit with a scalable communication network and interface. *Science*, *345*(6197), 668–673.

Milborrow, S., & Nicolls, F. (2014). Active shape models with sift descriptors and mars. In *Visapp (2)* (pp. 380–387).

Mirmehdi, M., Xie, X., & Suri, J. (2008). *Handbook of texture analysis*. World

Scientific.

Mondloch, C. J., Le Grand, R., & Maurer, D. (2002). Configural face processing develops more slowly than featural face processing. *Perception*, *31*(5), 553–566.

Müller, M. (2007). Dynamic time warping. *Information retrieval for music and motion*, 69–84.

Neave, R. (1998). Age changes to the face in adulthood. *Craniofacial Identification in Forensic Medicine*, 215–231.

Nefian, A. V., & Hayes III, M. H. (1998). Hidden markov models for face recognition. *choice*, *1*, 6.

Ojala, T., Pietikainen, M., & Maenpaa, T. (2002). Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Transactions on pattern analysis and machine intelligence*, *24*(7), 971–987.

Park, U., Tong, Y., & Jain, A. K. (2010). Age-invariant face recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *32*(5), 947–954.

Pears, R., Widiputra, H., & Kasabov, N. (2013). Evolving integrated multi-model framework for on line multiple time series prediction. *Evolving Systems*, *4*(2), 99–117.

Phillips, P. J., Grother, P., Micheals, R., Blackburn, D. M., Tabassi, E., & Bone, M. (2003). Face recognition vendor test 2002. In *Analysis and modeling of faces and gestures, 2003. amfg 2003. ieee international workshop on* (p. 44).

Phillips, P. J., Moon, H., Rauss, P., & Rizvi, S. A. (1997). The feret evaluation methodology for face-recognition algorithms. In *Computer vision and pattern recognition, 1997. proceedings., 1997 ieee computer society conference on* (pp. 137–143).

Phillips, P. J., Moon, H., Rizvi, S. A., & Rauss, P. J. (2000). The feret evaluation methodology for face-recognition algorithms. *IEEE Transactions on pattern analysis and machine intelligence*, *22*(10), 1090–1104.

Phillips, P. J., Scruggs, W. T., O'Toole, A. J., Flynn, P. J., Bowyer, K. W., Schott, C. L., & Sharpe, M. (2010). Frvt 2006 and ice 2006 large-scale experimental results. *IEEE transactions on pattern analysis and machine intelligence*, *32*(5), 831–846.

Pishchulin, L., Gass, T., Dreuw, P., & Ney, H. (2012). Image warping for face recognition: From local optimality towards global optimization. *Pattern Recognition*, *45*(9), 3131–3140.

Pujol, F. A., & García, J. C. (2012). Computing the principal local binary patterns for face recognition using data mining tools. *Expert Systems with Applications*, *39*(8), 7165–7172.

Ramanathan, N., & Chellappa, R. (2006). Modeling age progression in young faces. In *Ieee computer society conference on computer vision and pattern recognition*

(Vol. 1, pp. 387–394).

Rao, C. R., & Kshirsagar, A. (1978). A semi-markovian model for predator-prey interactions. *Biometrics*, 611–619.

Rappaport, T. S., et al. (1996). *Wireless communications: principles and practice* (Vol. 2). Prentice Hall PTR New Jersey.

Rhodes, H. T. F. (1956). *Alphonse bertillon, father of scientific detection.* Abelard-Schuman.

Ricanek Jr, K., & Tesafaye, T. (2006). Morph: A longitudinal image database of normal adult age-progression. In *Automatic face and gesture recognition, 2006. fgr 2006. 7th international conference on* (pp. 341–345).

Rowland, D. A., & Perrett, D. I. (1995). Manipulating facial appearance through shape and color. *IEEE Computer Graphics and Applications*, *15*(5), 70–76.

Sai, P. K., Wang, J. G., & Teoh, E. K. (2015). Facial age range estimation with extreme learning machines. *Neurocomputing*, *149*(Part A), 364–372. Retrieved from `http://dx.doi.org/10.1016/j.neucom.2014.03.074` doi: 10.1016/j.neucom.2014.03.074

Sakoe, H., & Chiba, S. (1978). Dynamic programming algorithm optimization for spoken word recognition. *IEEE transactions on acoustics, speech, and signal processing*, *26*(1), 43–49.

Sakoe, H., Chiba, S., Waibel, A., & Lee, K. (1990). Dynamic programming algorithm optimization for spoken word recognition. *Readings in speech recognition*, *159*, 224.

Scandrett, C. M., Solomon, C. J., & Gibson, S. J. (2006). A person-specific, rigorous aging model of the human face. *Pattern Recognition Letters*, *27*(15), 1776–1787.

Schmid, H. (2013). Probabilistic part-ofispeech tagging using decision trees. In *New methods in language processing* (p. 154).

Scott, D. W. (2015). *Multivariate density estimation: theory, practice, and visualization.* John Wiley & Sons.

Seo, H., Song, Y., Kim, C., & Kim, H. (2016). Characteristics of korean childrens facial anthropometry evaluated by three-dimensional imaging. *Journal of the International Society for Respiratory Protection Vol*, *33*(1).

Shyama, V., & Linda, M. (2016). Hidden markovian model combined with dynamic mode decomposition for detecting deception in videos.

Sirovich, L., & Kirby, M. (1987). Low-dimensional procedure for the characterization of human faces. *Josa a*, *4*(3), 519–524.

Song, S., Miller, K. D., & Abbott, L. F. (2000). Competitive hebbian learning through spike-timing-dependent synaptic plasticity. *Nature neuroscience*, *3*(9), 919–926.

Spillantini, M. G., Schmidt, M. L., Lee, V. M.-Y., Trojanowski, J. Q., Jakes, R., & Goedert, M. (1997). α-synuclein in lewy bodies. *Nature*, *388*(6645), 839–840.

Suo, J., Zhu, S.-C., Shan, S., & Chen, X. (2010). A compositional and dynamic model for face aging. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *32*(3), 385–401.

Syambas, N. R., & Purwanto, U. H. (2012). Image processing and face detection analysis on face verification based on the age stages. In *7th international conference on telecommunication systems, services, and applications (tssa)* (pp. 289–293).

Tan, X., Chen, S., Zhou, Z.-H., & Liu, J. (2009). Face recognition under occlusions and variant expressions with partial similarity. *IEEE Transactions on information forensics and security*, *4*(2), 217–230.

Tan, X., Chen, S., Zhou, Z.-H., & Zhang, F. (2005). Recognizing partially occluded, expression variant faces from single training image per person with som and soft k-nn ensemble. *IEEE Transactions on Neural Networks*, *16*(4), 875–886.

Tang, X., & Wang, X. (2003). Face sketch synthesis and recognition. In *Computer vision, 2003. proceedings. ninth ieee international conference on* (pp. 687–694).

Tian, Y.-l., Kanade, T., & Cohn, J. F. (2000). Recognizing lower face action units for facial expression analysis. In *Automatic face and gesture recognition, 2000. proceedings. fourth ieee international conference on* (pp. 484–490).

Tu, E., Kasabov, N., Othman, M., Li, Y., Worner, S., Yang, J., & Jia, Z. (2014). Neucube (st) for spatio-temporal data predictive modelling with a case study on ecological data. In *2014 international joint conference on neural networks (ijcnn)* (pp. 638–645).

Uçar, A., Demir, Y., & Güzeliş, C. (2016). A new facial expression recognition based on curvelet transform and online sequential extreme learning machine initialized with spherical clustering. *Neural Computing and Applications*, *27*(1), 131–142.

Uiboupin, T., Rasti, P., Anbarjafari, G., & Demirel, H. (2016). Facial image super resolution using sparse representation for improving face recognition in surveillance monitoring. In *Signal processing and communication application conference (siu), 2016 24th* (pp. 437–440).

Vapnik, V. N., & Vapnik, V. (1998). *Statistical learning theory* (Vol. 1). Wiley New York.

Štruc, V., & Pavešić, N. (2010). The complete gabor-fisher classifier for robust face recognition. *EURASIP Advances in Signal Processing*, *2010*, 26.

Wang, X., & Tang, X. (2009). Face photo-sketch synthesis and recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *31*(11), 1955–1967.

Wegbreit, E., Weissman, A. B., Cushman, G. K., Puzia, M. E., Kim, K. L., Leibenluft, E., & Dickstein, D. P. (2015). Facial emotion recognition in childhood-onset bipolar i disorder: an evaluation of developmental differences between youths and adults. *Bipolar disorders*, *17*(5), 471–485.

Welch, L. R. (2003). Hidden markov models and the baum-welch algorithm. *IEEE*

Information Theory Society Newsletter, *53*(4), 10–13.

Widiputra, H. (2011). *Integrated multi-model framework for adaptive multiple time-series analysis and modelling* (Unpublished doctoral dissertation). Auckland University of Technology.

Widiputra, H., Pears, R., & Kasabov, N. (2011). Multiple time-series prediction through multiple time-series relationships profiling and clustered recurring trends. In *Pacific-asia conference on knowledge discovery and data mining* (pp. 161–172).

Widrow, B., & Stearns, S. D. (1985). Adaptive signal processing. *Englewood Cliffs, NJ, Prentice-Hall, Inc., 1985, 491 p., 1*.

Witten, I. H., & Frank, E. (2005). *Data mining: Practical machine learning tools and techniques*. Morgan Kaufmann.

Wolberg, G. (1990). *Digital image warping* (Vol. 10662). IEEE computer society press Los Alamitos, CA.

Wu, B., Hu, B.-G., & Ji, Q. (2016). A coupled hidden markov random field model for simultaneous face clustering and tracking in videos. *Pattern Recognition*.

Wu, T., Turaga, P., & Chellappa, R. (2012). Age estimation and face verification across aging using landmarks. *IEEE Transactions on Information Forensics and Security*, *7*(6), 1780–1788.

Yamada, T., Yamashita, K., Ishii, N., & Iwata, K. (2006). Text classification by combining different distance functions withweights. In *Seventh acis international conference on software engineering, artificial intelligence, networking, and parallel/distributed computing (snpd'06)* (pp. 85–90).

Yan, S., Zhou, X., Liu, M., Hasegawa-Johnson, M., & Huang, T. S. (2008). Regression from patch-kernel. In *Computer vision and pattern recognition, 2008. cvpr 2008. ieee conference on* (pp. 1–8).

Zhang, G., Huang, X., Li, S. Z., Wang, Y., & Wu, X. (2004). Boosting local binary pattern (lbp)-based face recognition. In *Advances in biometric person authentication* (pp. 179–186). Springer.