Age group classification based on facial images

Sai, Phyo Kyaw

2014

Sai, P. -K. (2014). Age group classification based on facial images. Master's thesis,
Nanyang Technological University, Singapore.

<https://hdl.handle.net/10356/59908>

<https://doi.org/10.32657/10356/59908>

Age Group Classification Based on Facial Images

Sai Phyo Kyaw

School of Electrical & Electronic Engineering

A thesis submitted to the Nanyang Technological University
in partial fulfilment of the requirement for the degree of
Master of Engineering

2014

Acknowledgement

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

I would also like to thank my co-supervisor Dr Wang Jiangang, currently a Research Scientist at the Institute for Infocomm Research (I2R), A*STAR. I greatly acknowledge him for his wise advice, ideas and his involvement in our research. Without his guidance and collaborations, this research work would not be possible.

I am forever indebted to my parents for providing me with love and encouragement to make this research work as painless as possible. I thank all staff members, friends and students in Computer Vision Lab at the School of Electrical and Electronic Engineering, Nanyang Technological University for their kind support and help.

Finally, I would like to acknowledge the School of Electrical and Electronic Engineering, Nanyang Technological University, for providing the award of a scholarship and the research facilities for my work in NTU.

Summary

There are a lot of possible real world applications where age estimation could be used. Age estimation from face images is a challenging problem since aging is a personalized process and it is also affected by many factors. Recently there is increasing interest in age estimation work but most researchers used publicly available aging face datasets most of which may not be suitable for age estimation work. First, some datasets contain face images of low quality; pose variation and extreme expression and they have limited number of images. Second, some datasets contain several images of the same person for an age. The condition makes age estimation performance on those datasets irrelevant to real world applications.

In this thesis, we proposed a framework for collecting a large set of aging face images over internet. A large aging face dataset where there are many higher quality images of different people for each age was collected. In our approach, the dataset was collected over internet using Microsoft Search Engine Bing API with age-related text query. By using different combinations of filter options of API, a large number of images could be collected and by looping through ages, images of different ages could be obtained. This automated method of age-labeled image collection is able to provide us the dataset required for research work with less resource of our own.

The second contribution is that we have evaluated the age estimation methods by using the existing and new datasets. For age estimation research work, performance improvement was emphasized since our initial accuracy was low. Our initial age estimation system was employed

with AdaBoost classifier where tree classifiers were used as weak learners taking Gabor features as inputs. Since tree classifier works on local points, AdaBoost acts like local feature selector in this method. Biologically Inspired Feature was introduced to improve the performance in our existing framework.

Then a new framework where Extreme Learning Machine classifier took output of dimension reduction methods like PCA on local features as inputs was introduced to get a much better performance. In this way, this new method worked on global of local features where the first method totally missed out global features. In our experiments, Local Gabor Binary Pattern was found out to be a better feature than Biologically Inspired Feature in the proposed new framework. Moreover, performance and learning speed were improved significantly higher compared to previous framework.

Though some performance improvement has been achieved in our age group classification work, aging feature extraction method still needs to be improved so that possible real world application could be implemented. For our work on four age group classifications, improvement of 40-50% accuracy to 70-80% accuracy was achieved on average overall performance from previous to current work. Individual class performance on middle age groups was still in an unsatisfactory state even though performance on younger and older age group was quite high. This result pointed to us that a good feature extraction method is still lacking in age estimation domain or there is no distinct feature present to differentiate middle-age groups.

Contents

Acknowledgement.....	i
Summary.....	ii
Contents.....	iv
List of Figures.....	ix
List of Tables.....	xii
List of Symbols and Abbreviations.....	xiii
1 Introduction	1
1.1 Motivation	1
1.2 Objectives	2
1.3 Original Contributions of the Thesis.....	4
1.4 Organization of the Thesis	5
2 A Literature Review	7
2.1 Introduction	7
2.2 Representation of Aging Features on Faces	8
2.2.1 Anthropometric Models	8
2.2.2 Active Appearance Models	9

2.2.3	Aging Pattern Subspace	10
2.2.4	Aging Manifold	12
2.2.5	Appearance Feature Models	14
2.3	Age Estimation Classifiers	15
2.4	Feature Extraction Methods	18
2.4.1	Gabor Filter	18
2.4.2	Local Binary Patterns (LBP)	20
2.4.3	Biologically Inspired Features (BIF)	21
2.4.4	Local Gabor Binary Patterns (LGBP)	23
2.5	Classification Algorithms	24
2.5.1	AdaBoost	24
2.5.2	Extreme Learning Machine (ELM)	26
2.6	Dimension Reduction Methods	29
2.6.1	Principal Component Analysis (PCA)	29
2.6.2	Linear discriminant analysis (LDA)	30
2.6.3	Locality Preserving Projections (LPP)	31
2.6.4	Orthogonal Locality Preserving Projections (OLPP)	32
2.7	Concluding remarks	33

3	A Novel Approach for Aging Face Image Database Collection	35
3.1	Introduction	35
3.2	Aging Database Collection with API Approach over Internet Images	35
3.2.1	Gathering Module	37
3.2.2	Extraction Module	41
3.2.3	Validation Module	44
3.3	Comparison of our collected dataset with existing databases	45
3.4	Concluding remarks	46
4	An Age Estimation Framework with AdaBoost Classifier	48
4.1	Introduction	48
4.2	Experimental Datasets.....	49
4.2.1	Dataset I.....	49
4.2.2	Dataset II	51
4.2.3	Dataset III.....	52
4.3	AdaBoost Classifier with Error Correction Output Codes.....	55
4.4	Experimental Results and Discussion	59
4.4.1	Result of dataset I	59
4.4.2	Result of dataset II	60

4.4.3	Result of dataset III.....	61
4.5	Concluding remarks	62
5	A Novel Age Estimation Framework with Extreme Learning Machines	63
5.1	Introduction	63
5.2	Extreme Learning Machine with Dimension Reduction Techniques	63
5.3	Experimental Results and Discussion	68
5.3.1	Experience on dataset I	68
5.3.2	Experience on dataset II	70
5.3.3	Experience on dataset III.....	74
5.4	A comparison study on all different methods on each dataset	76
5.4.1	Performance comparison on different methods on dataset I	76
5.4.2	Performance comparison on different methods on dataset II.....	77
5.4.3	Performance comparison on different methods on dataset III.....	78
5.5	Concluding remarks	79
6	Conclusion and Recommendations for Future Work.....	80
6.1	Conclusion.....	80
6.2	Recommendations for Future Work.....	82
6.2.1	Aging Feature Extraction, Algorithm and Dataset Collection	82

6.2.2	Age Range Estimation with Combination of Far-Distance Age Predictors.....	83
6.2.3	Smile and Get Your Age Prediction	84
Author's Publications		86
Bibliography		87

List of Figures

Figure 2-1 six ratios of measurements used to separate babies from adults (adopted from [20]) ..8	8
Figure 2-2 Samples of training image, shape model and intensity model (adopted from [22]) ..10	10
Figure 2-3 AGingpattErn Subspace (adopted from [25])11	11
Figure 2-4 Geometric interpretation of CEA (a) CEA perspective with the data normalization (b) Correlation measure of the objective function (adopted from [5]).....14	14
Figure 2-5 Illustration of the idea of locally adjusted robust regression (adopted from [51])17	17
Figure 2-6 (a) real and (b) imaginary components of a typical Gabor filter perspective view ...19	19
Figure 2-7 Local Binary Pattern (LBP) processing on image21	21
Figure 2-8 Processing steps in BIF (adapted from [41]).....22	22
Figure 2-9 How weak classifiers form strong classifier $H(x)$ on AdaBoost classification25	25
Figure 2-10 Structure of Extreme Learning Machine27	27
Figure 3-1 Collection of age-labeled images using API service.....36	36
Figure 3-2 Gathering module.....38	38
Figure 3-3 Extraction module41	41
Figure 3-4 Validation module44	44
Figure 3-5 Overview of our proposed image collection system.....47	47

Figure 4-1	Sample images of bad quality in FG-NET dataset.....	49
Figure 4-2	Sample images of dataset I A(child) B(teen) C(adult) D(senior adult).....	50
Figure 4-3	Samples images of dataset II	52
Figure 4-4	Samples images of dataset III	53
Figure 4-5	Framework with AdaBoost classifier using ECOC for multi-classification	56
Figure 4-6	Gabor response in 5 scales and 8 orientations	57
Figure 4-7	Methods of Biologically Inspired Features	58
Figure 4-8	(10) fold cross validation results of Gabor vs. Bio-Inspired Feature on dataset I	59
Figure 4-9	(10) fold cross validation results of Gabor vs. Bio-Inspired Feature on dataset II....	60
Figure 4-10	(10) fold cross validation results of Gabor vs. Bio-Inspired Feature on dataset III.	61
Figure 4-11	Performance of Gabor and BIF by AdaBoost on dataset III	62
Figure 5-1	Age estimation framework with Extreme Learning Machine	64
Figure 5-2	Local Gabor Binary Pattern processing.....	65
Figure 5-3	Performance comparison between LGBP and BIF with respective dimension reduction methods on dataset I	69
Figure 5-4	Individual class performance on BIF and LGBP on dataset I	70
Figure 5-5	Performance comparison between LGBP and BIF with respective dimension reduction methods on dataset II	71

Figure 5-6 Performance of BIF and LGBP with LDA dimension reduction method on dataset II	72
Figure 5-7 Individual class performance on BIF and LGBP on dataset II	73
Figure 5-8 Performance comparison between LGBP and BIF with respective dimension reduction methods on dataset III.....	74
Figure 5-9 Individual class performance on LGBP on dataset III	75
Figure 5-10 Performance comparison among combination of feature representations and classification algorithms on dataset I	76
Figure 5-11 Performance comparison among combination of feature representations and classification algorithms on dataset II.....	77
Figure 5-12 Performance comparison among combination of feature representations and classification algorithms on dataset III.....	78
Figure 6-1 Age estimation with category selection.....	82
Figure 6-2 Age classification framework with combination of far-distance age classifiers	84

List of Tables

Table 3-1 Comparison of our dataset with existing databases	45
Table 4-1 Age distribution of dataset I.....	50
Table 4-2 Age distribution of dataset II.....	51
Table 4-3 Age distribution of dataset III	53

List of Symbols and Abbreviations

x_i	Image instance
l_i	Image label
L	Label vector
x^g	Gabor kernel output of an image
$o_{u,v}(\cdot)$	Gabor output by filter of scale v and orientation u
$\phi_{u,v}(\cdot)$	Gabor kernel with orientation variable u and scale variable v
X	Aligned image space
X_g	Gabor space
x^b	Biologically Inspired Feature of an image
X_b	Biologically Inspired Feature space
$G_t(\cdot)$	Strong classifier at t iteration
$h_t(x_i)$	Weak learner at t iteration
α_t	Weight given to weak learner at t iteration
X_{lg}	Local Gabor Binary Pattern space
x'_i	Sample output obtained by a specified approach
w_i	Input weight vector
β_i	Weight vector connecting the i^{th} hidden node
H	Hidden layer output matrix
β	Output weight matrix
$X_{B+Method}$	Reduced dimension space on BIF feature by “Method” approach

$X_{LG+Method}$	Reduced dimension space on LGBP feature by “Method” approach
l'_i	Predicted label
L'	Predicted label vector
μ	Mean
σ	Standard deviation
API	Application programming interface
LBP	Local Binary Pattern
BIF	Biologically Inspired Feature
LGBP	Local Gabor Binary Pattern
ECOC	Error correction output codes
SVM	Support vector machine
ELM	Extreme Learning Machine
SLFN	Single-hidden layer feedforward neural networks
PCA	Principal component analysis
LDA	Linear discriminant analysis
LPP	Locality preserving projections
OLPP	Orthogonal locality preserving projections

Chapter 1

1 Introduction

1.1 Motivation

There are quite a lot of interesting areas where age estimation method could be applied. For instance, aging population statistic on customers of a certain product, service or program would be very useful information to be applied in Electronic Customer Relationship Management (ECRM)[1]. The companies can grow their business by making strategic plans to meet different consuming habit, preference & needs and by recommending marketing plans to reach products and services to the intended parties. The possible use of age estimation in vending machines of age-restricted products and access control to internet pages would bring Human Computer Interaction (HCI) up to the next level.

Though there is such potential demand for application usage, age estimation via facial images remains a difficult problem due to different rate of aging process contributed by not only genetic factors but also other factors like living style, health and place of stay[2]. In certain cases differences in appearance between adjacent age groups are negligible, causing difficulties in the process of age estimation[3]. Based on current research work on the area, it is still a long way to come at creditable level of performance.

Looking at the current aging research work, most of the datasets used are not suitable for aging problem. FG-NET[4] and Yamaha Gender and Age YGA[5] datasets are commonly employed in literature work. Low image quality, great variation of facial poses and

expressions and unbalanced distribution of images across ages and limited number of ages are shortcoming in FG-NET dataset. In YGA dataset, each of 1600 subjects provides 5 near frontal images at the same age. They are not good representations of aging dataset suitable to be used for mentioned applications. Most of the time, image of the interested party will not be available in training dataset. How could we collect an aging dataset with large number of individual face images in the same age? And finally, age estimation itself is a changing problem since aging process is personal and uncontrollable. Is it now possible for machines to predict the age of a person to a reasonable age range?

1.2 Objectives

Current work on age estimation is mostly performed on small datasets like FG-NET[4] where one person might have several photos across different ages or large dataset like Yamaha Gender and Age (YGA) database where one person may have several photos on the same age. Those dataset will not be suitable for research work to apply for interested applications mentioned before because the aging representations in those datasets are different from the real world scenario. For those applications, the person to be predicted for the age will be unknown to the system. In most of the interested applications, it is not to predict the age of the person from his personalized aging dataset. Since performance on face recognition domain is very high, frequent presence of the same subject at the same age would bring the performance up. Performance on such aging datasets would not reflect the same performance level on proposed applications in which it is most likely that there is no image of test subject in training phase. Morph dataset II[6] has quite a number of face images of different people

for ages older than 16 years. But the dataset is lack of younger face images and the image quality is not so good.

So it is crucial to collect an aging dataset with large number of individual face images representing the same age label. Internet becomes a perfect place to collect such a dataset with its rich contents. With current development in social media work, information available through API services provided by internet search engines could be very useful. It also makes it possible to collect relevant information easily without using much resource of our own. In this thesis, we propose to collect a good aging face dataset by using Microsoft Bing API service[7]. The objective is to collect a more representative database for facial age estimation.

Another objective of this thesis is to find new aging features and classifier to improve the performance of our current Error-correcting output codes (ECOC) age estimation framework [8]. In this thesis, we will try to come out with a new age estimation framework which would outperform the existing one.

1.3 Original Contributions of the Thesis

Due to lack of a good aging face dataset, we have proposed a method to collect one in this thesis. We need a lot of resources like web crawler, time to collect age labeled face images over internet if we implement web crawling of our own. Moreover, we need to filter out non-face images from age-labeled images collected. Microsoft Search Engine Bing[7] provides users access to information of 1000 images per request. The advantage by using Microsoft Search Engine Bing is that we can access to image information without using much resource and it provides filtering properties and it allows us easy to get rid of non-face images. In this way, age-labeled face images were collected for an aging face dataset using API service provided by Microsoft Search Engine Bing. The dataset containing more face images of different people on each age would be ideal data for age estimation research work for possible applications of interest because it reflects real world scenario where possible applications could be employed.

Moreover, performance improvement is achieved through introducing new feature representation into previous system and implementing new age estimation framework. Our first framework consists of Gabor feature extraction from face images, tree classification performed on feature dimension and boosting overall performance by AdaBoost method and ECOC to achieve to multi-classification ability[9]. Since tree classifiers directly take a set of Gabor features performed on each pixel of the image, the system was very sensitive to the alignment and intra-face variation within class. So Biologically Inspired Feature (BIF)[10], where standard deviation values are taken over Gabor feature blocks, was employed to the system and performance improvement was obtained. Since AdaBoost combined tree classifiers which performs on local features, overall performance depends on local features

only. It neglects global feature representation. So a new age estimation framework was implemented where input features were global features obtained with dimension reduction method like Principle Component Analysis (PCA)[11], Linear Discriminant Analysis (LDA)[12], Locality Preserving Projections (LPP)[13] and Orthogonal Locality Preserving Projections (OLPP)[14] methods and Extreme Learning Machine[15] was used as classifier. Significant performance improvement was achieved compared to previous framework and Local Gabor Binary Pattern (LGBP)[16] was found to outperform BIF feature in this framework.

1.4 Organization of the Thesis

We have organized the rest chapters as follows:

Chapter 2 reviews the background material related to our understanding and development on age estimation work. The chapter covers representations of aging features on facial images, age estimation classifiers and their advantages and disadvantages. Moreover, features extraction methods, classification algorithms and dimension reduction methods applied in our frameworks were discussed.

Chapter 3 describes our proposed methods to collect a large aging dataset so that quality research could be conducted on age group classification problem. We discussed how we have collected a good aging dataset using Microsoft Bing API services over internet.

Chapter 4 includes discussion on our existing frameworks of age estimation, AdaBoost with Error Correction Output Codes and how performance improvement was achieved by employing Biologically Inspired Feature representation.

Chapter 5 covers introducing of new framework, Extreme Learning Machine with dimension reduction techniques and how better performance was achieved through introducing Local Gabor Binary Pattern to our framework.

Chapter 6 presents the conclusion of the thesis and suggestions for future research work in the area.

Chapter 2

2 A Literature Review

2.1 Introduction

Age progression displayed on faces is uncontrollable, individual and time dependent. Such special characteristics of aging variation cannot be captured somewhat accurately due to the prolific and diversified information conveyed by the human faces.

Even though aging process is personal and different from one another, people grow old with some common changes[17]. Two stages of gradual changes can be observed in human whole life[18]. During early aging of young people, the most prominent change is the geometrical change in their facial heads. When it comes to adult aging, the change in craniofacial growth becomes negligible and the skin changes (wrinkles) become perceptible. First, aging feature representation methods used in age estimation work will be mentioned and their advantages and disadvantages will be discussed. Then age estimation classifiers will be surveyed and their performance will be discussed. Furthermore, feature extraction techniques, classification methods and dimension reduction approaches applied in our age estimation frameworks will be briefly described.

2.2 Representation of Aging Features on Faces

2.2.1 Anthropometric Models

An extensive study of face anthropometry can be seen on Farkas' work[19]. Facial anthropometry is the study of measuring feature position in size and proposition. Face anthropometry were represented with propositions obtained from measurements among 57 facial landmarks over ages. Kwon and Lobo[20] made use of six different ratios of respective facial measures shown in **Figure 2-1** to distinguish babies from young adults and seniors.

Figure 2-1 six ratios of measurements used to separate babies from adults (adopted from [20])

Eight ratios of facial distance was proposed for age progression model in young age by Ramanathan and Chellappa[21].

Anthropometry model for age estimation is only good for analyzing young facial images since there is no much change of face shapes in adulthood. Kwon and Lobo[20] applied wrinkles analysis on several regions where wrinkles are common in old ages so that seniors could be classified from babies and young adults.

2.2.2 Active Appearance Models

The active appearance models are successfully applied for face encoding. It is a statistical face model proposed initially in [22] for coding facial images. A statistical shape model and an intensity model are trained separately over a given set of face images through the principal component analysis. A structural shape, points and intensity shape, the patch can be obtained by training on a number of labeled images.

Each image is featured with a set of landmark anthropometric points as showed in **Figure 2-2**. During training, AMM is performed on a set of images. The shape feature vectors, the appearance feature vectors and the shape-appearance feature vectors are obtained through PCA by eigen-decomposition. In this way, the shapes and appearances of the images are described in feature vectors[23].

Using active appearance model, Lanitis et al. [24] proposed a quadratic aging function, $age = \mathbf{f}(\mathbf{b})$, to estimate ages of face images. In the aging function, age is the actual age label on training phase and expected age of face image on testing phase, $\mathbf{b}$ is a vector containing 50 raw model parameters learned from active appearance model and $\mathbf{f}$ is the quadratic aging

function. The aging function finds the link between the age of individuals and AMM representation of the face images.

Figure 2-2 Samples of training image, shape model and intensity model (adopted from [22])

2.2.3 Aging Pattern Subspace

The aging progress is gradual and uncontrollable. Different people age in different ways. The aging patterns are temporal data. Considering those facts, an automatic age estimation method named AGES (AGingpattErn Subspace) was proposed Geng et al. [25]. The basic idea is to model the aging pattern, the sequence of personal face images sorted in the temporal order, by constructing a representative subspace. The test image is reconstructed into the aging pattern subspace and the position of the face image which yields lowest reconstruction error is taken as the predicted age.

Figure 2-3 AGingpattErn Subspace (adopted from [25])

Without taking each aging face image separately, a temporal sequence of aging face images from the same person could be employed all together to form the aging pattern model. If the face images of a person are available for all ages, the corresponding aging pattern is taken as a complete aging pattern. Otherwise it is considered as an incomplete aging pattern. There are two stages: the learning stage and the age estimation stage in AGES method. In the learning stage, aging pattern subspace is obtained by PCA technique. Initial values for some missing images in aging pattern would be filled by taking average between available images. The proposed EM-like iterative learning method is also used to minimize the reconstruction error

characterized by the difference between the available age images and the reconstructed face images. The mean face and eigenface of the covariance matrix of all face images could be computed. The face could be reconstructed from obtained eigenvectors and mean face. The iteration continues until the reconstruction error becomes acceptable.

In the age estimation stage, the test face is placed at every possible position in the aging pattern, respective mean face and eigenface are calculated for reconstruction and the position, at which the minimum reconstruction error occurs, is chosen as the predicted age label. In the aging pattern subspace approach, 200 AMM[22] features were driven from face images but Lanitis et al. [26] used 50 AMM features. However, the main difference of the AGES method is using personal face images at different ages altogether to represent the aging pattern.

2.2.4 Aging Manifold

Unlike training for a specific aging trend for the person, many face images of individuals at different ages are trained to obtain a common aging pattern. The aging information could be represented by a common feature from different faces of the same age. In this assumption, each individual could contribute many face images at a single age or at different ages. It becomes easier to do collection of aging database. The age manifold captures the common aging pattern by manifold embedding techniques. It is trained with many faces images at each age to obtain low dimensional aging trend.

The aging face image space could be represented by the set of facial images $X = \{x_i\}_{i=1}^n \in R^D$ of m ages with image dimension of $D = w \times h$ for n people. The labeling is represented by $L = \{l_i : l_i \in N\}_{i=1}^n$ for n people. The objective of aging manifold is to

obtain a low dimensional subspace representing intrinsic data information and geometric structures. So aging subspace could be represented as $Y = \{y_i\}_{i=1}^n \in R^D$ with $d \ll D$. Linear or non-linear methods could be used in the mapping function P which acquires the manifold space from the image space. The mapping function must be satisfied with the condition: $Y = P(X, L)$. After obtaining $D \times d$ matrix $\mathbf{P}$ satisfying $\mathbf{Y} = \mathbf{P}^T \mathbf{X}$ in a linear way [27] or a non-linear way [28, 29] during training, the testing images are transformed into $\mathbf{Y}'$ by multiplying $\mathbf{P}^T$ and $\mathbf{T}$ where $\mathbf{Y} = [\mathbf{y}_1 \mathbf{y}_2 \dots \mathbf{y}_n]$, $\mathbf{T} = [\mathbf{t}_1 \mathbf{t}_2 \dots \mathbf{t}_n]$, $\mathbf{X} = [\mathbf{x}_1 \mathbf{x}_2 \dots \mathbf{x}_n]$ and $\mathbf{P} = [\mathbf{p}_1 \mathbf{p}_2 \dots \mathbf{p}_d]$.

A statistical age estimation with manifold embedding was proposed by Scherbaum et al. [30]. The 3D morphable model information was embedded into a low dimensional subspace by PCA method. Prediction of age was performed by finding the curved trajectory of isosurface projected by support vector regression.

To connect the manifold learning with subspace, Cai et al. [14] proposed Orthogonal Locality Preserving Projection (OLPP), an orthogonal basis function obtained from Locality Preserving Projection (LPP) [13, 31] formulation which finds essential manifold structures on local neighbourhood distances. OLPP method was found to represent the aging features better than traditional PCA method in their experiment. Moreover, Conformal Embedding Analysis, shown in **Figure 2-4**, is employed to get improvement in aging estimation by Fu et al. [5].

The embedding low dimensional spaces by LPP or CEA could be obtained in supervised style not like in traditional PCA where only unsupervised way is possible.

Liwen Lu and Yap Peng Tan [32] proposed Ordinary Preserving Manifold Analysis for human age and head pose estimation. Their approaches : ordinary preserving LDA and

ordinary preserving MFA were devised from Linear Discriminant analysis (LDA) and MFA[27].

Figure 2-4 Geometric interpretation of CEA (a) CEA perspective with the data normalization (b) Correlation measure of the objective function (adopted from [5])

2.2.5 Appearance Feature Models

Both global and local features representation have been used in age estimation work in the literature. Hayashi et al. [33, 34] used features on texture (wrinkle) as well as shape (geometry) in aging and gender modeling work. Fukai et al.[35] extracted Fast Fourier Transform (FFT) features from facial images and selected local features using Genetic Algorithm (GA). Texture information may not be captured well for such method. Local Binary Pattern (LBP)[35] was used as texture features in automatic age classification work[36]. Gabor features are used in age classification task[37] and was found to be a better texture feature than LBP in their experiment. Spatially Flexible Patch (SFP) was employed as aging appearance descriptors by Yan et al. [38] to manage misalignment, occlusion and

pose variation of small changes effectively. Hierarchical face model with topology, geometry, photometry and configuration was proposed by Suo et al.[39] based on their earlier work, multi-resolution dynamic model for face aging simulation[40]. In this model, appearance features were extracted on different parts with respective filters at all levels. Motivated by object recognition in cortex[41, 42], Biologically Inspired Features (BIF) was proposed to use in aging estimation work by Guo et al. [10]. The features processing contains two layers; the first layer S1 obtained by Gabor features extraction methods on face images and the second layer C1 obtained by standard deviation values on the patches of the first layer. This kind of features could capture better pattern on face images of small translation, rotation and scale changes. Han et al. [43] proposed BIF features extracted from individual facial components for their hierarchical age estimation approach.

2.3 Age Estimation Classifiers

Both classification and regression can be found in age estimation work. Taking each age or age group as a class label, age estimation becomes a classification problem. If aging is taken as a continuous process of changes from young to old, it is a regression problem.

Artificial Neural Network (ANN) was first employed for 4-class age-group classification work by Kanno et al.[44]. Comparison between different classifiers on automatic age estimation was found on Lanitis' work [26]. Performance was evaluated on nearest neighbor, ANN, quadratic function classifier (regression) and Self-Organizing Map (SOM) methods with a small dataset of 400 images containing age range of 0 to 35 years. Regression method of quadratic function was found to perform better than ANN and SOM but not as good as nearest neighbor. Clustering with unsupervised method was suggested to apply on facial

images before actual classification or regression process. Expectation-Maximization (EM) algorithm was employed with Gaussian models in low dimensional 2DLDA + LDA features by Ueki et al. [45] for 11 group classification of WIT-DB database. The test image was filled into each Gaussian models and prediction was obtained by comparing the likelihoods.

A multiple linear regression method was used for age estimation in Conformal Embedding Analysis manifold by Fu et al.[5]. Improvement on performance was achieved significantly over existing methods in their experience on YGA database. Semi-Definite Programming (SDP) was introduced to formulate age estimation problem by Yan et al.[46]. The regression function was trained from uncertain non-negative levels. It performed better than quadratic function and multi-layer perceptron on YGA database. Even though SDP formulation gives better performance, it is a computationally expansive method. It is difficult for SDP to realize in a very large training database. An EM method was presented for optimization again in aging regression work by Yan et al. [47]. Age regression was performed on Spatially Flexible Patch (SFP) representation modeled by Gaussian Mixture Model (GMM) and improved by mean of patch-kernel measurement [48]. Image Based Regression (IBR) was employed on Haar-like features selected by boosting method to estimate the age in Zhou's work [49]. IBR improved significantly in computational efficiency in their 5 fold cross validation test on FG-NET database. Age group specific Linear Regression (ALR), Multi-Layers Perceptron (MLP), Support Vector Regression (SVR) and logistic regression were evaluated for age estimation on FG-NET and their private dataset by Suo et al. [39]. MLP was found to be the best performer in their experience. Robust Multi-Instance Regression (RMIR) was employed to deal with wrong labeling in a large web collected database for proposed universal age

estimator [50]. By mean of web mining, training could be updated and the system achieved reasonable performance.

Performance of classification and regression on aging problem was compared by Guo et al. [51]. Support Vector Machine (SVM) was selected as a representative classifier and Support Vector Regression (SVR) as a representative regressor. SVM achieved much higher accuracy than SVR on YGA database but SVR achieved higher performance than SVM on FG-NET database.

Figure 2-5 Illustration of the idea of locally adjusted robust regression (adopted from [51])

Locally Adjusted Robust Regressor (LARR)[51] was proposed combining SVM and SVR so that good features of each method could be achieved as shown in **Figure 2-5**. First, SVR was applied to predict all possible ranges of ages from the domain and SVM was applied on a small range group for classification to get better results. LARR method gave higher accuracy than individual method.

Hu Han et al. [43] proposed a hierarchical age estimation method in which each facial component is first classified into one of four disjoint age groups using a binary decision tree based on SVM(SVM-BDT) and then to the final age prediction by a separate SVM age regressor.

SVM was used to classify young and adult groups and SVR, SVO, Linear and PLS Regression were used to estimate age in respective young and adult groups in hierarchical age estimation [23]. Geng et al.[52] propose an algorithm named IIS-LLD for learning from the label distributions, which is an iterative optimization process based on the maximum entropy model to overcome the lack of training data in aging face domain.

2.4 Feature Extraction Methods

2.4.1 Gabor Filter

Gabor filtering is the linear filtering method invented by Dennis Gabor to use for edge detection in image processing. Gabor filters employed frequency and orientation representations which resembles the visual system of human and it makes the method the appropriate approach for texture representation and discrimination. The visual cortex of mammalian brains could be constructed by Gabor functions [53]. So image analysis with Gabor wavelets is quite comparable to perception in human visual system [54]. In the spatial domain, Gabor filter could be formulated as a Gaussian kernel function driven by a sinusoidal plane wave. Gabor filters are similar to one another and one wavelet could be formed into another wavelet by means of dilation and rotation.

Gabor filter's impulse response is defined by a harmonic function multiplied by Gaussian function. If it is in Fourier transform, multiplication will become convolution due to the convolution theorem.

The filter output value has a real and an imaginary component where they are orthogonal to one another. So the Gabor output could be represented as complex value or individual value as follow:

Complex

$$g(x, y; \lambda, \theta, \psi, \sigma, \gamma) = \exp\left(-\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2}\right) \exp\left(i\left(2\pi \frac{x'}{\lambda} + \psi\right)\right) \quad (2.1)$$

Real part:

$$g(x, y; \lambda, \theta, \psi, \sigma, \gamma) = \exp\left(-\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2}\right) \cos\left(2\pi \frac{x'}{\lambda} + \psi\right) \quad (2.2)$$

Imaginary part:

$$g(x, y; \lambda, \theta, \psi, \sigma, \gamma) = \exp\left(-\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2}\right) \sin\left(2\pi \frac{x'}{\lambda} + \psi\right) \quad (2.3)$$

where $x' = x \cos \theta + y \sin \theta$ and $y' = -x \sin \theta + y \cos \theta$.

Figure 2-6 (a) real and (b) imaginary components of a typical Gabor filter perspective view

In this equation, λ represents the wavelength of the sinusoidal factor, θ represents the orientation of the normal to the parallel stripes of a Gabor function, ψ is the phase offset, σ is the sigma of the Gaussian envelope and γ is the spatial aspect ratio, and specifies the ellipticity of the support of the Gabor function.

Gabor filters are the same as Gabor wavelets since they come up with different scales and rotations. However expansion is not desirable for Gabor wavelets due to the nature of time-consuming computational requirement. It is a common practice to create a Gabor filter bank preprocessed with various scales and orientations to reduce the computational cost in processing[55]. The Gabor space is obtained by the convolution of the filter values and the signal. This process functions quite similar to that of the primary visual cortex[56]. Jones and Palmer showed that the real part of the complex Gabor function could represent well to the receptive field weight functions working in simple cells in a cat's striate cortex[57].

They are widely used in image processing, computer vision, neuroscience and psychophysics. The Gabor space is very popular features used in image processing applications such as optical character recognition, iris recognition and fingerprint recognition.

2.4.2 Local Binary Patterns (LBP)

Local Binary Patterns (LBP) introduced in 1994, was another popular image feature for classification in computer vision [58, 59]. LBP feature is found to be powerful in texture classification. Moreover Histogram of Oriented Gradient (HOG) classifier modeled on LBP features gives the best classifier for classification of humans[60].

Figure 2-7 Local Binary Pattern (LBP) processing on image

The basic LBP feature vector is processed in the following way. First, the image is subdivided into cells (e.g. 16 x 16 pixels for each cell). In each cell, compare the processed pixel with its neighboring pixels along the specified direction (e.g. clockwise or counter-clockwise direction.) The representation value is processed as 0 where the neighboring pixel value is less than the center pixel value else as 1. In this way, binary number is obtained in digits limited by neighboring pixel numbers. The histogram is calculated on the number value over the cell. Normally the histogram is normalized to avoid big numbers. Finally normalized histograms of all cells are combined in a row or column to represent the LBP vector.

2.4.3 Biologically Inspired Features (BIF)

The features are obtained by combining the response of local edge-detectors that are slightly tolerant in position and scale over neighboring positions and multiple orientations which resemble complex cells in primary visual cortex. The features are more flexible than template-based approaches because small distortions of the input are allowed; they are more selective than histogram-based descriptors as local feature geometry is preserved.[41, 61]

The following section explains how BIF processing works as shown in **Figure 2-8**:

S1: The input image is transformed into Gabor maps by applying with multi-scale and multi-orientation Gabor filters. Only odd number pixel sizes are considered for the Gabor filter scale sizes. Then output maps are arranged in b bands which group Gabor maps with two subsequent scales in all orientations.

Figure 2-8 Processing steps in BIF (adapted from [41])

C1: For each band, the map is sub-sampled by taking the maximum over a grid with size of even number between two subsequent odd numbers used as pixel size for scaling in each

Gabor map. Then the maximum from the same spatial neighborhood over the two scales inside the same band are taken as (C1).

During training, K patches $P_{i=1,\dots,K}$ of various sizes $n_i \times n_i$ are extracted for all four orientations (thus containing $n_i \times n_i \times 4$ elements) at random from the (C1) maps from all training images.

S2: Prototype or centers of S2 are computed as the Gaussian-like response to all image patches X (at all positions). The responses are obtained by mean of the radial basis function (RBF): $Y = \exp(-\gamma \|X - P_i\|^2)$ and each patch P learned during training for each band independently.

C2: Corresponding to a particular patch P_i , the maximums over all positions and scales are computed for each S2 map and the shift- and scale-invariant C2 features are obtained for $i = 1 \dots K$.

2.4.4 Local Gabor Binary Patterns (LGBP)

In this approach, the face image is modeled as a histogram sequence by concatenating the LBP histograms of all the local regions obtained from all the local Gabor Magnitude Pictures. The effectiveness of the LGBP comes from several aspects including the multi-resolution and multi-orientation Gabor decomposition, the Local Binary Pattern, and the local spatial histogram modelling. LGBP actually consists of many pieces of histogram corresponding to different face components at different scales and orientations. [16] The face image is modeled as a “histogram sequence” by the following procedure:

-
- (1) The input face image is normalized and transformed to obtain multiple Gabor Magnitude Pictures (GMPs) by applying multi-scale and multi-orientation Gabor filters.
 - (2) Each Gabor Magnitude Picture is further divided into non-overlapping rectangle regions with specific size.
 - (3) Each region is applied with LBP and histogram is computed for each region.
 - (4) The LGBP histograms on all regions are concatenated to form the final histogram sequence as the model of the face.

2.5 Classification Algorithms

2.5.1 AdaBoost

Adaptive boosting mostly known as AdaBoost, is a popular machine learning algorithm proposed by Yoav Fred and Robert Schapire.[62] It is normally used together with other classifiers called weak learners to gain a better overall performance. AdaBoost is adaptive in a way that consecutive classifiers are employed to emphasize on those instances wrongly classified by previous classifiers. Since the classifier pays more attention to misclassified instances, noisy data and outliers are problems to AdaBoost. However it is found to be less susceptible to over-fitting problem than most of classification algorithms.[63]

AdaBoost could be processed as the following steps in most of boosting methods.

- given training set $(x_1, y_1), \dots, (x_m, y_m)$
- $y_i \in [-1, +1]$ correct label of instance $x_i \in X$

- for $t = 1, \dots, T$:
 - construct distribution D_t on $[1 \dots m]$
 - find weak classifier
 - $h_t : X \rightarrow [-1, +1]$ with small error ϵ_t on D_t
- output final classifier H_{final}

Figure 2-9 How weak classifiers form strong classifier $H(x)$ on AdaBoost classification

Distribution $D_t(i)$ could be updated as $\frac{D_t(i)}{Z_t}$ where Z_t is normalization constant. So $D_1(i) = 1/m$ so that $\sum_i D_t(i)$ will be kept as 1.

- Given D_t and h_t :

$$\begin{aligned}
D_{t+1}(i) &= \frac{D_t(i)}{Z_t} \times \begin{cases} e^{-\alpha_t} & \text{if } y_i = h_t(x_i) \\ e^{\alpha_t} & \text{if } y_i \neq h_t(x_i) \end{cases} \\
&= \frac{D_t(i) \exp(-\alpha_t y_i h_t(x_i))}{Z_t}
\end{aligned} \tag{2.4}$$

α_t is chosen as $\frac{1}{2} \ln \left(\frac{1-\epsilon_t}{\epsilon_t} \right)$ where $\epsilon_t = \sum_{i: y_i \neq h_t(x_i)} D_t(i)$. In this way, larger weights will be assigned to the instances where prediction is incorrect and the succeeding weak classifier will take serious consideration on those instances.

2.5.2 Extreme Learning Machine (ELM)

Extreme Learning Machine (ELM)[64] is a simple and efficient learning algorithm in single-hidden layer feedforward neural networks (SLFNs). ELM has several interesting and significant features over traditional popular gradient-based learning algorithms like feedforward neural networks.

Not like other feedforward network family, training could be done by ELM in an extremely short period. It can overcome the speed barrier faced by classic learning algorithms. Moreover it provides better generalization performance over the gradient-based training methods. The ELM tends to overcome several issues like local minima, improper learning rate and over-fitting problems straightforward. But traditional classical gradient-based learning algorithms may need to use additional methods like early stopping or weight decay to overcome such issues[64].

The ELM is a much simpler learning algorithm than most learning algorithms for feedforward neural networks. The present form of ELM algorithm is still only valid for

single-hidden layer feedforward networks (SLFNs). However, the fact that SLFNs can approximate any continuous function and implement any classification application is provided in[65]. Thus, reasonably speaking the proposed ELM algorithm can be efficiently used in many applications.

In ELM, the input weights and hidden layer biases of SLFNs are first randomly chosen[65]. After the input weights and the hidden layer biases are assigned randomly, SLFNs can be simply considered as a linear system and the output weights (linking the hidden layer to the output layer) of the SLFNs can be analytically determined through simple generalized inverse operation of the hidden layer output matrices.

Figure 2-10 Structure of Extreme Learning Machine

For N arbitrary distinct samples $(x_i - t_i)$, where $x_i = [x_{i1}, x_{i2}, \dots, x_{in}]^T \in R^n$ and $t_i = [t_{i1}, t_{i2}, \dots, t_{im}]^T \in R^m$, and standard SLFNs with $\tilde{N}$ hidden nodes and activation function $g(x)$ are mathematically modeled as

$$\sum_{i=1}^{\tilde{N}} \beta_i g_i(x_j) = \sum_{i=1}^{\tilde{N}} \beta_i g(w_i \cdot x_j + b_i) = o_j, j = 1, 2, \dots, N \quad (2.5)$$

where $w_i = [w_{i1}, w_{i2}, \dots, w_{in}]^T$ is the weight vector connecting the i^{th} hidden node and the input nodes, $\beta_i = [\beta_{i1}, \beta_{i2}, \dots, \beta_{im}]^T$ is the weight vector connecting the i^{th} hidden node. $w_i \cdot x_j$ denotes the inner product of w_i and x_j . The output nodes are chosen linear in the approach.

That standard SLFNs with $\tilde{N}$ hidden nodes with activation function $g(x)$ can approximate these N samples with zeros error means that $\sum_{j=1}^{\tilde{N}} ||o_j - t_j|| = 0$, i.e. there exist β_i , w_i and b_i such that

$$\sum_{i=1}^{\tilde{N}} \beta_i g(w_i \cdot x_j + b_i) = t_j \quad j = 1, 2, \dots, N \quad (2.6)$$

The above N equations can be written compactly as

$$H\beta = T \quad (2.7)$$

where

$$H(w_1, \dots, w_{\tilde{N}}, b_1, \dots, b_{\tilde{N}}, x_1, \dots, x_N) = \begin{bmatrix} g(w_1 \cdot x_1 + b_1) & \cdots & g(w_{\tilde{N}} \cdot x_1 + b_{\tilde{N}}) \\ \vdots & \ddots & \vdots \\ g(w_1 \cdot x_N + b_1) & \cdots & g(w_{\tilde{N}} \cdot x_N + b_{\tilde{N}}) \end{bmatrix}_{N \times \tilde{N}}$$

$$\beta = \begin{bmatrix} \beta_1^T \\ \vdots \\ \beta_{\tilde{N}}^T \end{bmatrix}_{\tilde{N} \times m}$$

$$T = \begin{bmatrix} t_1^T \\ \vdots \\ t_N^T \end{bmatrix}_{N \times m}$$

H is called the hidden layer output matrix of the neural network; the i^{th} column of H is the i^{th} hidden node output with respect to inputs $x_1, x_2, \dots, x_N$. The output weight β could be computed as $\beta = H^+T$ where H^+ is the pseudoinverse of hidden layer output matrix H .

ELM can be used efficiently in many applications. However the universal approximation capability and the performance of sparse high-dimensional applications are still open interesting aspects in ELM.

2.6 Dimension Reduction Methods

2.6.1 Principal Component Analysis (PCA)

Principal component analysis (PCA) is a mathematical formulation where an orthogonal transformation is used so that a set of observations of possibly correlated variables can be represented by a set of values of uncorrelated variables called principal components. Since some variables can be insignificant and lost in the data representation, it is normally lesser number in principal components than the number of original variables. This transformation is processed in such a way that the first principal component provides the largest possible variance and each subsequent component produces the highest variance possible under the constraint of being orthogonal to the previous components[66]. PCA can be mathematically computed by mean of eigenvalue decomposition of a data covariance matrix or singular value

decomposition of a data matrix after the data matrix of individual attribute are normalized to the center[67].

Find the embedding that maximizes the projected variance.

$$p = \arg \max_{||p||=1} p^T S p \quad (2.8)$$

where S is the scatter matrix

$$S = \sum_{i=1}^n (x_i - \bar{x})(x_i - \bar{x})^T$$

and $\bar{x}$ is the mean vector of $\{x_i\}_{i=1}^n$.

2.6.2 Linear discriminant analysis (LDA)

Linear discriminant analysis (LDA) or Fisher's linear discriminant performs in a way relatively close to principal component analysis (PCA) and factor analysis where linear combinations of variables are generated differently to represent the data in the best form.[68] PCA neglects any difference between classes and oppositely factor analysis obtains the feature combinations driven on differences rather than similarities. LDA explicitly considers the difference between the classes of data. Discriminant analysis differs from factor analysis in the sense that it is not an interdependence technique where a distinction between independent variables and dependent or criterion variables is considered. Continuous independent variables and a categorical dependent variable are considered in discriminant analysis[69]. The LDA method takes a fundamental assumption that the independent variables are normally distributed.[70]

The within-class scatter matrix is as follows

$$S_w = \sum_{j=1}^c \sum_{i=1}^{N_j} (x_i^j - \mu_j)(x_i^j - \mu_j)^T \quad (2.9)$$

The between class matrix is

$$S_b = \sum_{j=1}^c (\mu_j - \mu)(\mu_j - \mu)^T \quad (2.10)$$

where $w = \arg \max_w \frac{w^T S_b w}{w^T S_w w}$.

2.6.3 Locality Preserving Projections (LPP)

Locality Preserving Projection (LPP) is linear projective maps where a variation problem is solved in a way that the neighborhood structure of the data set is optimally taken into account. LPP objects to find an embedding manifold that preserves local information, and obtains a subspace that describes the essential manifold structure to the best. Locality Preserving Projections are achieved by getting the optimal linear approximations to the eigen functions of the Laplace Beltrami operator on the manifold so that the high dimensional data could be represented on a low dimensional manifold embedded in the ambient space.[31]

Though it is linear computation, LPP represent the data as nonlinear techniques such as Laplacian Eigenmaps[71] or Locally Linear Embedding[28] in most properties.

Find the embedding that preserves essential manifold structure by measuring the local neighborhood distance information.

$$p = \arg \min_{p^T X D X^T p = 1} \sum_{i=1}^n \sum_{j=1}^n (p^T x_i - p^T x_j)^2 s_{ij} \quad (2.11)$$

where

$$s_{ij} = \begin{cases} e^{(-||x_i - x_j||^2 / t)}, & ||x_i - x_j||^2 < \varepsilon \\ 0 & \text{otherwise} \end{cases},$$

$$s_{ij} = \begin{cases} e^{(-||x_i - x_j||^2 / t)}, & \text{when } \mathbf{x}_i \text{ and } \mathbf{x}_j \text{ are } k \text{ nearest neighbors of each other} \\ 0 & \text{otherwise} \end{cases}$$

$$S(i, j) = s_{ij}, D(i, j) = \sum_j s_{ji} \text{ and } L = D - S.$$

2.6.4 Orthogonal Locality Preserving Projections (OLPP)

In Locality Preserving Projections (LPP), their components are not orthogonal to one another and so it faces difficulty in rebuilding the data. Orthogonal Locality Preserving Projection method overcomes the difficulty by employing orthogonal basis functions. Thus it is more powerful than LPP in locality preserving strength. In other way, OLPP is likely to provide better discriminating power than LPP. It is to construct an adjacency graph reserving the geometry of the examined manifold and the class relationship between various points. The projection is finally a preserved graph structure which models the Orthogonal Laplacian face[72].

Orthogonal basis functions are derived based on LPP to maximize discrimination power for embedding.

$$A^{(i-1)} = [p_1 \dots p_{i-1}] \quad (2.12)$$

$$B^{(i-1)} = [A^{(i-1)}]^T (XDX^T)^{-1} A^{(i-1)} \quad (2.13)$$

p_1 is obtained by finding eigenvector of $Q = (XDX^T)^{-1}XDX^T$ associated with the smallest eigenvalue.

p_i is calculated as the eigenvector of $M = Q - (XDX^T)^{-1}A^{(i-1)}[B^{(i-1)}]^{-1}[A^{(i-1)}]^T$ associated with the smallest eigenvalue.[73]

2.7 Concluding remarks

Facial anthropometry model [20, 21] is totally not suitable for adult faces. Facial measurements are taken from 2D images, the face post become very sensitive issue for facial anthropometry model. Compared to the anthropometry model based approaches, the AAMs can be applied for young ages but also for old ages in general since the approach deals with both shape and texture information in image representation. For AGES method, there must have enough images of the same person at different ages to model the aging pattern. It is not an easy to collect large database having closed imaging conditions. Another problem is that AMM face features could not represent facial wrinkles well for senior adult people. Texture pattern is missing in its intensity encoding. Intensities in spatial neighborhood must be considered to obtained local texture information. Compared to the aging pattern subspace, the age manifold does not require a specific aging pattern learning approach for each individual. All images of different individuals across all ages could be employed to learn for aging

estimation. However the size of training images should be large enough so that the embedded manifold with statistical sufficiency could be achieved for the aging manifold methods. Texture features are better than other extraction methods since they consider neighborhood intensities in their processing.

It is difficult to take age estimation as classification or regression problem. The performance would depend on the database and the respective algorithms. However hybrid approach of classification and regression was found to get higher performance than individual method in age estimation work.

Moreover, feature extraction methods and classification algorithms which were to be applied in our work were discussed briefly. Efficiency of Extreme Learning Machine will be a good factor when generalization of the problem is to be determined.

Chapter 3

3 A Novel Approach for Aging Face Image Database Collection

3.1 Introduction

Lack of a good public aging dataset makes difficult for a quality research work on age estimation. Due to large multiclass nature of human ages, it is difficult to collect the aging database large enough to do quality research work on the problem. Aging database of high quality and quantity had been collected by mean of Application Programming Interface (API) services provided by Microsoft Search Engine Bing. Automatic collection of age-specified query images was implemented using C# programming interface with Microsoft BING API services. It is a good way to collect aging images in huge volume since internet hosts a large number of images with age labeling. It serves as an alternative platform to gather face images for age classification.

3.2 Aging Database Collection with API Approach over Internet Images

Motivated by works of web image mining towards universal age estimator[50] and web image mining age estimation framework[74], we tried to collect aging face dataset with Microsoft Programming Application Interface (API) services provided by Microsoft Search Engine Bing. It is not easy to implement a powerful web crawler which could provide image

search capability with required text match. Search engines employ web crawl feature to gather information from various websites over internet and fetch those information to us. With advance in social media network, we could access to a large amount of information through API services provided by social media service providers. Like traditional web user interface search engine, service providers will return the information of the related search queries to the users through supported protocols by respective API. The Internet search engine chosen for this project is the Microsoft Bing API[7] due to its advantages such as the most synchronized API service among service providers[75], no limitation of send queries and the rich documentation available on the internet.

Figure 3-1 Collection of age-labeled images using API service

Age-related text queries were sent to Bing service provider through API mean and information of matched image results were received and saved on the system. Then some filtering processing was done on acquired data and selected images were downloaded to the

system. Then manual validation were done on the downloaded images and labeled accordingly to their expression, gender and so on. This proposed approach was implemented in .NET platform with three main modules: namely Gathering Module, Extraction Module and Validation Module, to achieve web image mining for aging human facial images.

3.2.1 Gathering Module

This first module is responsible for communicating with BING engine to requesting age related query and to gather returned image search responses. To communicate with BING, protocol of SOAP based approach was used since it is the most sophisticated approach and it gets rid of the limitations of HTTP GET request of XML protocol.

In this module, an automated mechanism was implemented to perform the sending of the age related text queries to the Bing API service. It had the ability to send a list of age related request like “xx years old” and “xx birthday” to search for facial images from one to eighty years of age. There were 14 of such predefined text request statements for each age. This module also gathered all the return responses from the API from each search query request and classified the responses based on the age related text and will save them on the system. This gathering module had capability of error handling in a case of communication breakdowns between API and the module.

Figure 3-2 Gathering module

Due to limitation set by BING, a maximum of 1,000 results per query are allowed to access through API service provided. However, BING provides filter properties so that we could select desired category of image search response. The following are some of the filter properties provided by BING:

Size : <Small / Medium / Large>

Aspect: < Square/ Wide / Tall>

Face: <Face / Portrait>

So we could send request statement 12 times for each query by arranging different combination of three filter properties mentioned above. In this way we could access request results up to 12,000 per query. Moreover, we could remove most of images which do not contain human face by using face property in filter options since age related text statement

may be relevant to image labels given to birthday cakes or the age of the buildings, wines, animals and so on. This function performed a noise filtering that allowed only the desired human faces images to be return. Hence, by making use of this property a face detection function is not required. And age related query statement was cascated by the text string of “NOT singer NOT actor NOT actress NOT football NOT celebrity” so that photos of famous people would be eleminated. Since news agencies and people tend to share about life of famous people, repitative appearance of their same photos and large presence of their photos are common for the same query statement. Our ideal aging face database covers a large number of face images taken from different people but not from the same person.

Finally the gathered image informantion would be written and saved as binary data file inside repective age folder and the file would be named acccording to the text query statement for subsequent process. The flow of Gathering Module could be discribed as the following flow chart diagram.

Flow chart of Gathering Module

3.2.2 Extraction Module

Extraction module went into respective age folder and read through one binary file after another created by gathering module to perform an automated downloading of desired images from the Internet. Since there are 14 age-related query statements for each age, information of the same image might be gathered a few times. This module had the capability to prevent downloading of duplicated copies of the same image from the Internet. The basic concept is to compare all the Media URLs of downloaded images with URL of image to be extracted and not to download any image having the same media URL to the system again.

Figure 3-3 Extraction module

This mechanism was also served as an additional layer to perform a robust noise filtering to reduce undesired images from being downloaded. This filtering could be achieved by searching for some predefined text strings like age x , age_ x , x -year-old and requested text

query statement inside image title, URL and media URL indices. Only when the search string was found in those indices, the module performed downloading of the image from Internet. It is the observed fact that the human age information is naturally included in the image such as Media URL or Tags or Title for those human facial images. Motivated by this observation, the component will be added to this implementation to build a robust noise filter. This noise filtering was able to help the system to prevent any further noisy image from being extracted into the system, to reduce succeeding manual validation effort and also to provide label information for the image.

Extraction module will offer the various features:

- Automated checking for duplication copies of images before extraction from the Internet to prevent a duplication of images in database
- Automated extractions of aging human facial images over the Internet to reduce human effort
- Ability to perform robust noise filtering by using keywords attached to image information such as Media URL or Image Title.

The following is the flow chart for Extraction Module:

Flow chart of Extraction Module

3.2.3 Validation Module

The purpose of Validation Module is to perform a process of labeling of all images and validate downloaded images are desired aging human facial images.

Figure 3-4 Validation module

This module accessed each dataset of different search queries performed on each age and displayed all the images in a graphic user interface (GUI) for human validation. The process is to ensure face in the image is in the range of requested age and the information represents the image since web crawling may get image information through nearby text. In addition, the validation module allowed the users to further classify and categorize all the images according to the different categories such facial expression, gender, skin color, with face accessories and so on. However this manual approach to labeling and categorizing become a enormous task for a huge database.

This approach allows the applications to perform collection of large dataset of human facial images over all ages with minimum human manual work. Such aging face dataset would be useful for training or learning for an age estimation system. **Figure 3-5** describes the overview of the approach taken to develop this project and flow of the process and the key modules presented.

3.3 Comparison of our collected dataset with existing databases

The following table shows quantitative comparison of our approach and current existing databases. Table includes average number of images, μ and standard deviation, σ within specified ranges of ages. The mentioned ranges were selected at intervals where the number of the images is significantly changed in Morph II database, the biggest among existing databases.

Table 3-1 Comparison of our dataset with existing databases

		Age range			
		0-15	16-55	56-65	66-77
FG_NET	μ	35.93	10.45	0.70	0.20
	σ	5.87	9.98	0.10	0.57
PAL	μ	0	8.00	5.70	12.50
	σ	0	8.83	6.18	2.89
MORPH II	μ	0	1354.55	118.5	6.23
	σ	0	478.15	93.96	4.99
Our approach	μ	($\approx$)860.00	($\approx$)1316.33	($\approx$)877.00	($\approx$)949.33
	σ	($\approx$)228.68	($\approx$)288.30	($\approx$)224.00	($\approx$)204.11

As shown in Table 3-1, the average number of images of each age is quite low in FG_NET and PAL. There is nothing or negligible number of images in some age ranges. For Morph II,

significant number of images is in the range of age 16-55 with average of 1354 images, there is no image below the age of 15 and beyond age 55, the number of images in each age is reduced significantly and high standard deviation values compared to average values are reflected.

For our approach, not all the ages are considered for comparison. The ages we used in our experimental dataset were considered in calculation such as age 3, 5, 8, 10 and 13 for range 0-15, age 23, 25, 28, 30, 33, 43, 45, 48, 50 and 53 for range 16-55, age 63 and 65 for range 56-65 and 68, 70 and 73 for range 66-77. From the huge amount of images, only images with good pose, reasonable quality and low level of expression are taken for comparison. According to statistics, our approach is capable of providing the number of image over individual ages good enough to do decent research work in age estimation.

3.4 Concluding remarks

We have proposed an approach to collect aging face images of interest using freely available services provided over internet. It should be noticed that the proposed approach can be extend to other applications easily. From the database collected, we created a new dataset with quality images to do initial experiment. However the results obtained were found not to be satisfactory even the image normalized manually. Feature detection on face and computing resource are not good enough to do large scale computing on age estimation domain. However, the collected aging face database will be useful in near future due to increasing interest in the age estimation area. Automatic labelling or categorizing is a necessary step to materialize the idea.

Figure 3-5 Overview of our proposed image collection system

Chapter 4

4 An Age Estimation Framework with AdaBoost Classifier

4.1 Introduction

It is difficult to achieve the system which estimates an actual age from facial images and most applications do not need to have an exact or small-range age estimation performance. Classification on a small number of age groups was emphasized. Experience was done on three datasets for four age groups. In dataset I, images are classified into children, teen, adult and senior adult and in dataset II, images are classified into class I (0-19), class II (20-39), class III (40-59) and class IV (60-79). Then in final dataset III, images are grouped into four according to their age 16, 26, 36 and 46.

First, AdaBoost method combined with Error Correction Output Codes (ECOC)[9] was implemented for age group classification.[8] AdaBoost method tries to achieve the strong classifier from a set of weak classifiers where tree classifier was used in implementation. Since AdaBoost is binary classifier, it combines with ECOC method to provide multi-classification capability. The features extracted are Gabor filter output (8 orientations and 5 scales) and Biologically Inspired Features (BIF)[10].

4.2 Experimental Datasets

4.2.1 Dataset I

The dataset I consists of aging facial images from two public aging datasets; FG-NET[4] and PAL. FG-NET aging database was collected by European Working Group of Face and Gesture Recognition Research Network. Most of the FG-NET age labels are in the younger range and less number of images in old age range. The oldest age of the dataset is 69 years old. Some facial images of the FE-NET database are bad in quality. Some images have photo defects and marks on face, some images are blurring and some have extreme facial expressions and poses of different angles as shown in **Figure 4-1**. Such kind of bad images were left out for research work.

Figure 4-1 Sample images of bad quality in FG-NET dataset

The PAL database was created at University of Michigan by Meredith Minear and Denise Park. The PAL database does not contain face images of children and young teenagers. The stating age in the data base is 18 years old and oldest age is 93 years old. The quality and pose of face in images are good. Some of bad images and facial expression and pose of extreme were left out from FE-NET database. The dataset has four age groups namely child, teen, adult and senior adult. Good images from FG-NET were grouped to from child and teenage group and adult group. Divert age range from PAL dataset were selected to form

senior adult and added some to adult group to have reasonable amount of data in each group. Images are grouped into four categories as shown in **Table 4-1**.

Table 4-1 Age distribution of dataset I

Group	Age range	Number of images
Child	less than 10	96
Teen	11 to 19	71
Adult	20 to 60	116
Senior Adult	60 & above	100

Sample images of dataset I could be seen on **Figure 4-2**.

Figure 4-2 Sample images of dataset I A(child) B(teen) C(adult) D(senior adult)

4.2.2 Dataset II

The dataset II was created with the images collected over internet by the method which uses Microsoft BING Image Search Application Programming Interface (API). Querying with different age-labeled text string on API services, age-labeled images were collected and saved with relevant age-labeled information. The details of how image collection was done over internet were explained on section 3.2. From collected large pool of image data, image of good quality were selected from certain age and the information was verified manually. This dataset was created with facial images of good quality and where there is no so much variation on facial pose and not much serious condition of extreme expression for research work on facial aging of human. The age distribution and sample images of dataset II could be seen on the following:

Table 4-2 Age distribution of dataset II

Group	Age	Images / Age	Total Images
I	3	50	235
	5	50	
	8	44	
	10	51	
	13	40	
II	23	49	251
	25	50	
	28	49	
	30	57	
	33	46	

II	43	64	252
	45	52	
	48	39	
	50	51	
	53	46	
IV	63	35	225
	65	52	
	68	46	
	70	49	
	73	43

Figure 4-3 Samples images of dataset II

4.2.3 Dataset III

Dataset III was created with images from the Morph database II[6], a largest collection publicly available aging face database. It has facial images of male and female starting from 16 to 77 years of age. It contains 55,134 images of 13,000 individuals collected over four

years. Since the images are taken in real-world conditions, it is good for research work aimed for real-world applications. Male African images of age 16, 26, 36 and 46 were selected and used for this dataset with each group containing 250 individual images each. The reason for testing this dataset is to judge how good the performance on group of 10-year gap on the same profile is. The sample images and age distribution of dataset III can be seen in **Figure 4-4** and **Table 3**.

Figure 4-4 Samples images of dataset III

Table 4-3 Age distribution of dataset III

Group	Age	Number of images
I	16	250
II	26	250
III	36	250
IV	46	250

It is not necessary to obtain the precise estimates of the actual age for some applications[76], we believe that four age-group estimation is good enough to explore and discover underlying issues in age estimation domain. The images in all three datasets are aligned with eye corner points captured manually and cropped to the image size of 65 x 75.

4.3 AdaBoost Classifier with Error Correction Output Codes

Boosting is an iteration process and it can be expressed as Equation (3.1).

$$C_t(x_i) = \alpha_1 h_1(x_i) + \alpha_2 h_2(x_i) + \dots + \alpha_t h_t(x_i) \quad (3.1)$$

Each iteration t , C_t will select weak learner h_t which yields the lowest penalty given the current set of weights. A two-node tree classifier as weak learner h_t will perform on each of input dimension space at t iteration. So AdaBoost classifier also acted like dimension reduction method to the input high dimensional space in our framework. ECOC enabled AdaBoost classifier perform binary into multi-class classification[9]. It made multi-class problem into a combination of binary classifiers. 4 x 7 ECOC matrix was used for the experience as follow:

$$\begin{bmatrix} 1 & -1 & -1 & -1 & 1 & 1 & 1 \\ -1 & 1 & -1 & -1 & 1 & -1 & -1 \\ -1 & -1 & 1 & -1 & -1 & -1 & 1 \\ -1 & -1 & -1 & 1 & -1 & 1 & -1 \end{bmatrix}$$

It needed one AdaBoost classifier for one code word. Each column defined a binary partition of 4classes over data from which a binary classifier is trained. Each row represented the code word for the respective class. So a total of 7 strong classifiers were trained to accommodate code word of 7 digits. For example, the first classifiers will try to classify class 1 from the rest, the fifth one as class 1 & 2 against class 3 & 4 and the last one as class 1 & 3 against class 2 & 4 using (4 x 7) ECOC matrix mentioned above. Finally the sample was classified into the class whose code word is the closest to the combination of predicted output sequence. There will be some kind of error correction capacity with using ECOC since it is possible to

recover the code word from the prediction sequence even certain classifier gets its part wrong.

There will be three modules in the framework of age group classification. **Figure 4-5** illustrates the diagram of the proposed framework.

Figure 4-5 Framework with AdaBoost classifier using ECOC for multi-classification

Suppose the image space I is represented by a set of aligned faces images $X = \{x_i\}_{i=1}^n \in R^{D_1}$ of m subjects in the order of subject ages. A label set $L = \{l_i : l_i \in N\}_{i=1}^n$ associated with the images provides the age group labeling. After features extraction process, the Gabor output space $X_g = \{x_i^g\}_{i=1}^n \in R^{D_2}$ or Biologically Inspired Feature (BIF) space $X_b = \{x_i^b\}_{i=1}^n \in R^{D_3}$ will be obtained and will be fed into AdaBoost Classifier for classification.

Figure 4-6 Gabor response in 5 scales and 8 orientations

Since Gabor features were extracted from the image at five scales and eight orientations as shown in **Figure 4-6**, the Gabor wavelet representation of the image x^g could be processed as in equation (3.2).

$$x^g = [o_{u,v}(x, y) : u \in \{0, \dots, 7\}, v \in \{0, \dots, 4\}] \quad (3.2)$$

where $o_{u,v}(x, y) = x_i(x, y) * \varphi_{u,v}(x, y)$ and $\varphi_{u,v}(x, y)$ is Gabor kernel; u and v represent the orientation and scale of Gabor filter. Gabor features are extracted from image of 65 by 75, Gabor output dimension will have 65x75x40 (195000). BIF extraction technique was not implemented as in [10]. But the idea behind is the same where standard deviation value of predefined blocks of Gabor image was taken. x^b was obtained by cascading the standard deviation values on the non-overlapping cells of the Gabor image, $o_{u,v}$ for all values of u and v .

Figure 4-7 Methods of Biologically Inspired Features

As shown in **Figure 4-7**, Gabor processing is performed on S1 layer and standard deviation are tabulated on block size of Gabor images on C1 layer. BIF feature is the cascading of those standard deviation values obtained.

AdaBoost classifier was boosted by combining of weak learners using tree classifiers. Since tree classifier took a few nodes as inputs, AdaBoost acted like selecting local features on which the combination of all selected tree classifiers gave the best performance at respective iteration. AdaBoost in this implementation would be sensitive towards alignment of the images since feature dimension were related to pixel location on the image.

4.4 Experimental Results and Discussion

The performance of Gabor features and BIF features was evaluated on AdaBoost framework with all the three datasets. A clear improvement on performance was found with Biologically Inspired Feature on all datasets. 10 fold cross validation where one fold is reserved for testing while all the other nine folds are used for training.

4.4.1 Result of dataset I

Dataset I has four age groups naming Child, Teen, Adult and Senior Adult. It was prepared from combination of FG-NET and PAL. Results obtained on 10 cross-validation fold can be seen in **Figure 4-8**.

Figure 4-8 (10) fold cross validation results of Gabor vs. Bio-Inspired Feature on dataset I

9 out of 10 folds, BIF feature provided higher accuracy than Gabor feature. Only on one fold, Gabor performed slightly better than BIF feature. On average, accuracy was improved on BIF feature by about 13%.

4.4.2 Result of dataset II

Dataset II was obtained from images collected over internets with API approach. It too has four groups and each contains 5 ages. Performance comparison of Gabor and BIF on dataset II is shown in **Figure 4-9**.

Figure 4-9 (10) fold cross validation results of Gabor vs. Bio-Inspired Feature on dataset II

Results clearly show that BIF feature provides better accuracy than Gabor feature. In all 10 folds, BIF feature was on top of Gabor feature. The system provided accuracy of 55% on BIF, which is 10% higher than that of Gabor.

4.4.3 Result of dataset III

Dataset III was created with images of age 16, 26, 36 and 46 from Morph database II.

Performance comparison of Gabor and BIF on dataset II is shown in **Figure 4-10**.

Figure 4-10 (10) fold cross validation results of Gabor vs. Bio-Inspired Feature on dataset III

Similar to the results of dataset I and II, BIF feature outperformed Gabor feature in all folds and obtained average mean of 50% which is also 13% higher than that of Gabor.

Performance analysis was considered up to class level as shown in **Figure 4-11**. Results of 10 fold cross validation was compared and presented. Pink dotted line represents the overall performance accuracy. Color bars below the pink line represent true positive values of representative classes and color bars above the line represent false negative values. As we could see from **Figure 4-11**, it is clear that the performance was improved clearly not only on overall level but also on individual class level.

Figure 4-11 Performance of Gabor and BIF by AdaBoost on dataset III

4.5 Concluding remarks

BIF feature which represents statistical value over Gabor feature block might have some tolerance towards alignment and orientation and might have a better representation of aging texture information. It is obvious that BIF feature yielded better performance than normal Gabor feature. Accuracy was improved overall as well as on individual classes. However, this method only performed on local features and global features were neglected.

Chapter 5

5 A Novel Age Estimation Framework with Extreme Learning Machines

5.1 Introduction

Previously AdaBoost classifier used tree classifiers as weak learners. The tree classifier with two nodes selected three feature points from high dimensional feature space. So AdaBoost classifier performed classification based on local feature points selected by weak learners. It never made use of its global features. To overcome this limitation, Extreme Learning Machine (ELM) approach was adopted.

We will discuss how we implemented a new framework which gave better accuracy than our previous framework and performance among difference feature representation methods.

5.2 Extreme Learning Machine with Dimension Reduction Techniques

In this experience, global features were extracted by means of dimension reduction techniques and classification work was done on reduced feature space. Moreover Extreme Learning Machine (ELM)[15] is a simple and efficient learning algorithm driven from single-hidden layer feedforward neural networks (SLFNs). ELM provides several interesting and significant features over traditional popular gradient-based learning. Not like other feedforward network family, training in ELM will last only an extremely short period.

Moreover it provides better generalization performance over the gradient-based training methods.

Figure 5-1 Age estimation framework with Extreme Learning Machine

Since the input weights and hidden biases are randomly chosen in ELM, learning time traditionally spent in tuning these parameters is irrelevant here. Random assigned computational nodes in the hidden layer are mostly likely to be independent of the training data. It makes ELM provides better performance on the functional approximation. ELM appears to be suitable in applications which request fast prediction and response capability.[77]

In this framework as shown in **Figure 5-1** , first different features were extracted from face images manually aligned at eye corner points. Experience was carried out to compare the results on features obtained by two extraction methods. Since Biologically Inspired Features tested on previous framework provided improvement on the system, the method will be

employed again and compared against Local Gabor Binary Pattern.[16] Four dimension reduction approaches will perform on the features extracted by those two methods.

Both features extraction techniques: Biologically Inspired Features (BIF) and Local Gabor Binary Pattern (LGBP), made use of Gabor features extracted with Gabor filter band of 5 scales and 8 orientations. Local Gabor Binary Pattern was implemented as in shown in **Figure 5-2**. It was first proposed for face recognition domain. All 40 Gabor images generated by Gabor filter bank are divided into processing blocks one after another. LBP histogram was computed on the block and cascaded into LBP vector. Since the extracted feature dimension was too high in both Biologically Inspired Features and Local Gabor Binary Pattern, it was not proper to do classification with limited number of images in high dimensional space.

Figure 5-2 Local Gabor Binary Pattern processing

Then dimension reduction was performed on the extracted feature space to get low dimension space from high dimension space. Different feature dimension reduction methods like Principle Component Analysis (PCA), Linear Discriminant Analysis (LDA), Locality

Preserving Projections (LPP) and Orthogonal Locality Preserving Projections (OLPP) were used and compared. In the previous AdaBoost framework, most discriminant local features were selected by tree classifiers and AdaBoost classifier boosts the performance by combining the weak learners. So the features used in that framework were just local. In this ELM framework for age estimation, global representation of features was extracted to low dimensional space and classification was performed on the extracted feature space with ELM classifier to predict the age group. Classification work could be formulated as the following equations.

$$\begin{bmatrix} g(w_1 \cdot x'_1) & \cdots & g(w_N \cdot x'_1) \\ \vdots & \ddots & \vdots \\ g(w_1 \cdot x'_N) & \cdots & g(w_N \cdot x'_N) \end{bmatrix} \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_N \end{bmatrix} = \begin{bmatrix} l_1 \\ \vdots \\ l_N \end{bmatrix} \quad (5.1)$$

$$H\beta = L \quad (5.2)$$

where x'_i is i^{th} sample value obtained by a specified method, w_i is the input weight vector connecting the i^{th} hidden node and the input node, β_i is the weight vector connecting the i^{th} hidden node and l_i is the label.

The output weight β could be obtained by the equation (5.3).

$$\beta = H^+L \quad (5.3)$$

where H^+ is the pseudoinverse of hidden layer output matrix H .

Suppose the image space I is represented by a set of aligned faces images $X = \{x_i\}_{i=1}^n \in R^{D_1}$ of n subjects. A label set $L = \{l_i : l_i \in N\}_{i=1}^n$ associated with n images provides a total of N age group labeling. Then Biologically Inspired Feature (BIF) space $X_b = \{x_i^b\}_{i=1}^n \in R^{D_3}$ and Local Gabor Binary Pattern (LGBP) space $X_{lg} = \{x_i^{lg}\}_{i=1}^n \in R^{D_4}$ or

will be obtained. PCA, LDA, LPP and OLPP methods are used to reduce their high dimensional space and to obtain some sense of their global representation. So a total of 8 combination methods will be obtained as (BIF+PCA) representation $X_{B+PCA} = \{x_i^1\}_{i=1}^n \in R^{D_5}$, (BIF+LDA) representation $X_{B+LDA} = \{x_i^2\}_{i=1}^n \in R^{D_6}$, (BIF+LPP) representation $X_{B+LPP} = \{x_i^3\}_{i=1}^n \in R^{D_5}$, (BIF+OLPP) representation $X_{B+OLPP} = \{x_i^4\}_{i=1}^n \in R^{D_5}$, (LGBP+PCA) representation $X_{LG+PCA} = \{x_i^5\}_{i=1}^n \in R^{D_5}$, (LGBP+LDA) representation $X_{LG+LDA} = \{x_i^2\}_{i=1}^n \in R^{D_6}$, (LGBP+LPP) representation $X_{LG+LPP} = \{x_i^3\}_{i=1}^n \in R^{D_5}$ and (LGBP+OLPP) representation $X_{LG+OLPP} = \{x_i^4\}_{i=1}^n \in R^{D_5}$. Then X_{B+PCA} , X_{B+LDA} , X_{B+LPP} , X_{B+OLPP} , X_{LG+PCA} , X_{LG+LDA} , X_{LG+LPP} and $X_{LG+OLPP}$ will be fed into ELM for classification and their respective prediction output $L' = \{l'_i : l'_i \in N\}_{i=1}^n$ will be obtained.

Experience was done on two datasets as mentioned in dataset section. Since BIF and LGBP features are processed on blocks of Gabor output, the square block sizes which gave the best performance on BIF and LGBP was discovered from window size of 4x4 to 10x10 by 10 fold cross validation. Then their best performance results were compared and analyzed.

5.3 Experimental Results and Discussion

5.3.1 Experience on dataset I

On the first dataset of FGNET+PAL, it contains four age groups of child, teen, adult and senior adult. Overall performance as well as individual class performance was considered so that we could be able to decide kinds of feasible applications to work on.

Performance comparison was conducted on 10 fold cross validation where each group of images are equally divided into 10 folds and one fold is always reserved for testing and the other nine folds are used for training.

Overall performance comparison on all eight approaches was shown in **Figure 5-3**. From the results obtained, we could see that LGBP feature representation method provided better results than BIF representation method across all dimension reduction techniques even though BIF representation yielded better accuracy than LGBP representation on some folds. Unsupervised dimension reduction technique PCA performs reasonably well on Biologically Inspired Feature representations.

Figure 5-3 Performance comparison between LGBP and BIF with respective dimension reduction methods on dataset I

From the performance of different approaches, it is obvious that LGBP feature extraction method with LDA dimension reduction approach performed better than all other approaches. For individual class performance on LGBP + LDA approach, improvement was significant on all classes except on teen group from Biologically Inspired Feature to Local Gabor Binary Pattern. Child and senior adult groups performed quite well even on BIF feature domain.

However we could see unsatisfactory performance accuracy on teen group in the best approach as shown in **Figure 5-4**.

Figure 5-4 Individual class performance on BIF and LGBP on dataset I

Even overall classification performance of four classes is quite high and close to 80%, the accuracy on teen class could fall below 50% on most folds. The skin condition might be the factor which provided the best performance score on senior adults. Teen group suffered due to their similarity to both child and adult groups.

5.3.2 Experience on dataset II

In the second dataset collected over internet, it also contains four age groups of 0-19, 20-39, 39-60 and 61-80. Experience was done with 10 fold cross validation. The second dataset has 5 quite equally separated ages in each group. So the images in each age were equally distributed for 10 folds. Overall accuracy comparison of 10 fold cross validation on the second dataset was shown in **Figure 5-5**.

Figure 5-5 Performance comparison between LGBP and BIF with respective dimension reduction methods on dataset II

Similar to what we observed on the first dataset, LGBP feature representation method still performed better than BIF feature representation method. Moreover we could see the clear gap on results between supervised dimension reduction methods of LGBP and BIF across all folds. Images of better quality in dataset II might make LGBP representation provide more stable performance. Moreover, all of dimension reduction methods followed the similar trend of performance across the folds. They all performed well on certain folds and bad on other

folds respectively. LGBP + LDA approach is again top performer among all combinations.

There was lowest performance on Class III (40-59) for all methods.

Figure 5-6 Performance of BIF and LGBP with LDA dimension reduction method on dataset II

According to the figures mentioned above, LGBP method achieved higher individual performance than BIF method on class I, II and IV. The accuracy on class III was still poor and worsened even on LGBP method.

Figure 5-7 Individual class performance on BIF and LGBP on dataset II

The class I has the highest performance since faces of child and young teen of Class I are easier to differentiate from those of adults. The lowest performance on Class III reflects the fact that 40s and 50s are quite difficult to differentiate from 30s and 60s. Lack of good differentiating features in adult ages is obvious.

5.3.3 Experience on dataset III

Dataset III was created from Morph II database containing four age groups with 10-year gap i.e. 16, 26, 36 and 46. Experience was carried out to find out the performance accuracy on face images of 10-year gap ages.

Figure 5-8 Performance comparison between LGBP and BIF with respective dimension reduction methods on dataset III

Again in dataset III, LGBP performed better than BIF. But LGBP + LDA approach was no longer top performer. LGBP + OLPP came out as the best approach with accuracy of 68.5%. As individual class performance, it was the worst performance on age 36 group with lower

than 50% accuracy on most folds. Performance accuracy was reasonably good on other age groups.

Figure 5-9 Individual class performance on LGBP on dataset III

When the best performer LGBP + OLPP was analyzed, individual performance on class I and IV is reasonably good but performance on class II and III is unsatisfactory. The difference between 16 and 26 is more obvious than the difference between 36 and 46. This condition makes the group (16) obtain the highest accuracy whereas it makes the group (36) get the lowest.

5.4 A comparison study on all different methods on each dataset

5.4.1 Performance comparison on different methods on dataset I

From the starting framework (AdaBoost classifier with Gabor feature) to our current work, the performance improvement on dataset I could be seen as shown in **Figure 5-10**.

Figure 5-10 Performance comparison among combination of feature representations and classification algorithms on dataset I

Performance improvement was achieved through introducing new features as well as new framework. Our initial accuracy value was below 50% with AdaBoost + Gabor and we could

improve the accuracy up to 80% with ELM + LDA. We could see our new framework could provide us with a gain of 10% over previous framework. However new features employed did not introduce much improvement.

5.4.2 Performance comparison on different methods on dataset II

Figure 5-11 Performance comparison among combination of feature representations and classification algorithms on dataset II

In this dataset II, performance improvement was not as good as that of dataset I. However significant improvement could be observed step by step through each introduction of new approach in **Figure 5-11**. In each introduction, there was 8-10% performance improvement

achieved. As the result, initial 50% accuracy to final 75% accuracy was obtained through in this way.

5.4.3 Performance comparison on different methods on dataset III

Figure 5-12 Performance comparison among combination of feature representations and classification algorithms on dataset III

Comparison of results on this dataset III was presented in **Figure 5-12**. Previous AdaBoost framework obtained overall accuracy of 37% with Gabor feature. Then accuracy went up to 55% using BIF feature representation method. A new age estimation framework with ELM

was implemented and overall accuracy up to 60% was obtained using BIF feature and finally accuracy of 70% was achieved with LGBP feature method.

5.5 Concluding remarks

We have proposed a new framework for facial age estimation. By adopting ELM, improvements on accuracy were observed whenever new feature extraction methods or new framework was introduced. The overall performance of final stage was brought up to reasonably high, in-between groups still yielded low accuracy. New framework can manage to handle global features which might contribute to performance improvement. Moreover, ELM is extremely fast in learning process and it was a great help for our research work.

Chapter 6

6 Conclusion and Recommendations for Future Work

6.1 Conclusion

In this thesis, we highlighted most of existing aging databases used in age estimation research could not well reflect possible scenario in real world applications of interest. Results obtained on those databases could not be reliable for predicating the age from unknown faces in real applications. We have proposed a novel method to collect face images over internet for a better aging dataset to be used in facial age estimation. In our method, age-labeled face images were collected through API service provided by Microsoft Image Search Engine Bing with age-related text query. In this way, a large number of face images of different people were collected for each age and this has made the dataset a better representation of possible real world scenario where age estimation system might be applied. Comparing with the existing aging database, our method has provided a cheap approach to solve the problem in the existing face database, i.e. lack of the faces at different range.

Furthermore, we have developed some ways to improve performance on age group classification from previous to current stage. To increase accuracy, we introduced feature representation which has tolerance toward pose and alignment variation of a few degrees to previous system of AdaBoost classifier equipped with ECOC for multi-classification. 10-15% performance improvement was achieved depending on dataset by employing BIF feature in

previous system. Tree classifiers operating on local features are combined and boosted by AdaBoost methods. AdaBoost classifier used only local features which were selected by tree classifier during iteration process in this method. Even though ECOC method was used to achieve multi-classification capability, it might jeopardize the performance if the aging process happens to be continuous linear process since ECOC employed code word which is larger than the number of the classes.

Then a new age estimation framework was proposed with Extreme Learning Machine (ELM) as classifier taking output of dimension reduction methods. In this way, the new system will has ability to take global feature representation as inputs. Another advantage over previous system is that training process is really fast in ELM. Through experiments, ELM was found to be not only a good starter to understand the status quo of the working domain but also a good optimizer which provides reasonable generalization and reliable performance within a short time. Additional 10-15% performance improvement was obtained through this new framework. In this way, we can lift initial average overall accuracy of 40-50% on previous system to average overall accuracy of 70-80% on current system. Though we can achieve such a high overall performance for four age group classification, we found that the individual class performance on middle groups was still low. We need a breakthrough in aging feature extraction methods so that computer-based age estimation system could be employed to detect small age range difference. This could be our future work.

6.2 Recommendations for Future Work

6.2.1 Aging Feature Extraction, Algorithm and Dataset Collection

According to the results obtained, current aging features extracted are not distinct enough to do classification in middle age range. We need to work on aging feature extraction methods so that reasonable performance could be achieved on age estimation system. Since face variations among people of different expressions, races and gender are more dominate and obvious than face variation due to aging, we need to consider those facts seriously. Currently most of feature extraction methods employed are still affected by those changes. In other words, aging features extracted would be represented differently due to those facts or changes. So detection of such facts and changes should be considered in aging estimation algorithms as shown in **Figure 6-1**.

Figure 6-1 Age estimation with category selection

In this framework, the category of the face image will be first decided with category selector. So it will choose one of the different combinations of ethnicity, expression, gender and etc. Based on the result obtained, respective classifier will predict the age of the face image.

To do such a research, collection of aging face dataset with required information becomes important. We need a dataset where images of people with different expressions and with or without glasses should be collected. Moreover their ethnicity and gender information should be included. Images should be preferably taken at controlled environment.

6.2.2 Age Range Estimation with Combination of Far-Distance Age Predictors

Since face aging is personalized and gradual process of changes, it is not easy to detect age difference from face images in small age range. However we could start to notice difference when their age gap becomes wider. Based on the concept that it is easier to differentiate faces between wider gaps of ages, we would like to work on a new age range estimation method. In this method, there will be two steps: (1) a series of initial outputs will be predicted by initial classifiers, (2) final prediction result will be calculated by the decision making classifier performed on the series outputs obtained earlier.

Selection of age range for initial classifiers will be crucial in the approach so that good accuracy is obtained at initial stage. Assume employed classifier could provide a good performance on 10-year age gap. As shown in **Figure 6-2**, initial classifiers would be trained on respective 10-year age gap groups and the final classifier would only predict the age range based on the initial results.

Figure 6-2 Age classification framework with combination of far-distance age classifiers

6.2.3 Smile and Get Your Age Prediction

Smile is a way of greeting and a kind of showing friendliness. So we do it very often and we are good at it. It is more natural and easier to smile than to express other forms of emotions. However, the changes caused by face wrinkles are different on different age groups when expression of emotions is conveyed. So it would give us better age estimation results if face images at different expression stages will be classified with different classifiers. For example, the examinee will be requested to smile and images at benchmark action stages will be recorded. Then different classifiers trained on a specified expression stage will process on the face image at respective stage. Finally the final classifier will predict the age or age range from the results obtained by the different stages.

For this purpose, we need an aging face database in which smiling face images at different degrees are labeled and categorized accordingly.

Author's Publications

Conference Papers

- (1) Phyto-Kyaw Sai, Jian-Gang Wang and Eam-Khwang Teoh, "*Facial age estimation with Extreme Learning Machines*", Accepted in the 4th International Conference on Extreme Learning Machines (ELM2013), Beijing, China, October 15-17, 2013.
- (2) Phyto-Kyaw Sai, Jian-Gang Wang and Eam-Khwang Teoh, "*Web image mining for facial age estimation*", Accepted in the 9th International Conference on Information, Communications and Signal Processing (ICICS), Taiwan, Dec 10-13, 2013.

Journal Papers

- (3) Phyto-Kyaw Sai, Jian-Gang Wang and Eam-Khwang Teoh, "*Study on facial age estimation*", Submitted to Pattern Analysis and Applications in Dec 2013.
- (4) Phyto-Kyaw Sai, Jian-Gang Wang and Eam-Khwang Teoh, "*Facial age range estimation with Extreme Learning Machines*", Neurocomputing, special issue on ELM2013, accepted, March 2014.

Bibliography

- [1] R. S. Winer, "A Framework for Customer Relationship Management," *CALIFORNIA MANAGEMENT REVIEW*, vol. 43, pp. 89-105, 2001.
- [2] T. E. Seeman and E. Crimmins, "Social Environment Effects on Health and Aging," *Annals of the New York Academy of Sciences*, vol. 954, pp. 88-117, 2001.
- [3] A. Lanitis. (20 June 2012). *Facial Age Estimation*. Available: http://www.scholarpedia.org/article/Facial_Age_Estimation
- [4] (Oct 10 2011). *FG-NET Aging Database*. Available: <http://www.fgnet.rsunit.com/>
- [5] Y. Fu and T. S. Huang, "Human Age Estimation with Regression on Discriminative Aging Manifold," *IEEE Transactions on Multimedia*, vol. 10, pp. 578-584, 2008.
- [6] K. Ricanek and T. Tesafaye, "MORPH: A Longitudinal Image Database of Normal Adult Age-progression," in *Proceedings of the 7th International Conference on Automatic Face and Gesture Recognition*, 2006, pp. 341-345.
- [7] Microsoft. (Oct 29 2011). *Bing API*. Available: <http://msdn.microsoft.com/en-us/library/dd900818.aspx>
- [8] J.-G. Wang, W.-Y. Yau, and H. L. Wang, "Age Categorization via ECOC with Fused Gabor and LBP Features " in *Proceedings of Workshop on Applications of Computer Vision*, 2009, pp. 1-6.
- [9] T. G. Dietterich and G. Bakiri, "Solving Multiclass Learning Problems via Error Correcting Output Codes," *Journal of Artificial Intelligence Research*, pp. 263–286, 1995.
- [10] G. Guo, G. Mu, Y. Fu, and T. S. Huang, "Human Age Estimation using Bio-Inspired Features," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2009, pp. 112-119.
- [11] M. A. Turk and A. P. Pentland, "Face Recognition using Eigenfaces," in *Preceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 1991, pp. 586 - 591
- [12] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, "Eigenfaces vs. Fisherfaces: Recognition using Class Specific Linear Projection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, pp. 711 - 720 1997.
- [13] X. He, S. Yan, Y. Hu, P. Niyogi, and H.-J. Zhang, "Face Recognition using Laplacianfaces," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, pp. 328-340, 2005.
- [14] D. Cai, X. He, J. Han, and H.-J. Zhang, "Orthogonal Laplacianfaces for Face Recognition " *IEEE Transactions on Image Processing*, vol. 15, pp. 3608 - 3614 2006.
- [15] G. Huang, Q. Zhu, and C. Siew, "Extreme Learning Machine: Theory and Applications," *Neurocomputing*, vol. 70, pp. 489-501, 2006.

-
- [16] W. Zhang, S. Shan, W. Gao, X. Chen, and H. Zhang, "Local Gabor Binary Pattern Histogram Sequence (LGBPHS): A Novel Non-statistical Model for Face Representation and Recognition " in *Proceedings of the IEEE International Conference on Computer Vision ICCV*, 2005, pp. 786 - 791.
 - [17] Y. Fu, G. Guo, and T. S. Huang, "Age Synthesis and Estimation via Faces: A Survey," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, pp. 1955-1976, 2010.
 - [18] A. M. Albert, K. Ricanek Jr, and E. Patterson, "A Review of the Literature on the Aging Adult Skull and Face: Implications for Forensic Science Research and Applications," *Forensic Science International*, vol. 172, pp. 1-9, 2007.
 - [19] L. Farkas, *Anthropometry of the Head and Face*. New York: Raven Press, 1994.
 - [20] Y. H. Kwon and N. d. V. Lobo, "Age Classification from Facial Images," *Computer Vision and Image Understanding*, vol. 74, pp. 1-21, 1999.
 - [21] N. Ramanathan and R. Chellappa, "Modeling Age Progression in Young Faces," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2006, pp. 387-394.
 - [22] T. F. Cootes, G. J. Edwards, and C. J. Taylor, "Active Appearance Models," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, pp. 681-685, 2001.
 - [23] S. Kohli, S. Prakash, and P. Gupta, "Hierarchical Age Estimation with Dissimilarity-based Classification," *Neurocomputing*, vol. 120, pp. 164-176, 2013.
 - [24] A. Lanitis, C. J. Taylor, and T. F. Cootes, "Toward Automatic Simulation of Aging Effects on Face Images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, pp. 442-455, 2002.
 - [25] X. Geng, Z.-H. Zhou, and K. Smith-Miles, "Automatic Age Estimation Based on Facial Aging Patterns," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, pp. 2234-2240, 2007.
 - [26] A. Lanitis, C. Draganova, and C. Christodoulou, "Comparing Different Classifiers for Automatic Age Estimation," *IEEE Transactions on Systems, Man, and Cybernetics: Part B: Cybernetics*, vol. 34, pp. 621-628, 2004.
 - [27] S. Yan, D. Xu, B. Zhang, H.-J. Zhang, Q. Yang, and S. Lin, "Graph Embedding and Extensions: A General Framework for Dimensionality Reduction," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, pp. 40-51, 2007.
 - [28] S. T. Roweis and L. K. Saul, "Nonlinear Dimensionality Reduction by Locally Linear Embedding," *SCIENCE*, vol. 290, pp. 2323-2326, 2000.
 - [29] J. B. Tenenbaum, V. d. Silva, and J. C. Langford, "A Global Geometric Framework for Nonlinear Dimensionality Reduction," *SCIENCE*, vol. 290, pp. 2319-2323, 2000.
 - [30] K. Scherbaum, M. Sunkel, H. P. Seidel, and V. Blanz, "Prediction of Individual Non-Linear Aging Trajectories of Faces," *Computer Graphics Forum*, vol. 26, pp. 285-294, 2007.
-

-
- [31] X. He and P. Niyogi, "Locality Preserving Projections," in *Proceedings of Conference on Advances in Neural Information Processing System*, 2003.
 - [32] J. Lu and Y.-P. Tan, "Ordinary Preserving Manifold Analysis for Human Age and Head Pose Estimation," *IEEE Transactions on Human-Machine Systems*, vol. 43, pp. 249-258, 2013.
 - [33] J. Hayashi, M. Yasumoto, H. Ito, and H. Koshimizu, "Method for Estimating and Modeling Age and Gender using Facial Image Processing," in *Proceedings of the 7th International Conference on Virtual Systems and Multimedia*, 2001, pp. 439-448.
 - [34] J. Hayashi, M. Yasumoto, H. Ito, Y. Niwa, and H. Koshimizu, "Age and Gender Estimation from Facial Image Processing," in *Proceedings of the 41st SICE Annual Conference*, 2002, pp. 13-18.
 - [35] H. Fukai, H. Takimoto, Y. Mitsukura, and M. Fukumi, "Apparent Age Estimation System based on Age Perception," in *Proceedings of SICE Annual Conference*, 2007, pp. 2808-2812.
 - [36] A. Gunay and V. V. Nabiyev, "Automatic Age Classification with LBP," in *Proceedings of the 23rd International Symposium on Computer and Information Sciences* 2008, pp. 1-4.
 - [37] F. Gao and H. Ai, "Face Age Classification on Consumer Images with Gabor Feature and Fuzzy LDA Method," in *Proceedings of the 3rd International Conference on Advances in Biometrics*, 2009, pp. 132-141.
 - [38] S. Yan, M. Liu, and T. S. Huang, "Extracting Age Information from Local Spatially Flexible Patches," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, 2008, pp. 737-740.
 - [39] J. Suo, T. Wu, S. Zhu, S. Shan, X. Chen, and W. Gao, "Design Sparse Features for Age Estimation using Hierarchical Face Model," in *Proceedings of the 8th IEEE International Conference on Automatic Face & Gesture Recognition*, 2008, pp. 1-6.
 - [40] J. Suo, F. Min, S. Zhu, S. Shan, and X. Chen, "A Multi-Resolution Dynamic Model for Face Aging Simulation," in *Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition*, 2007, pp. 1-8.
 - [41] T. Serre, L. Wolf, S. Bileschi, M. Riesenhuber, and T. Poggio, "Robust Object Recognition with Cortex-Like Mechanisms," in *Proceedings of IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2007, pp. 411-426.
 - [42] M. Riesenhuber and T. Poggio, "Hierarchical Models of Object Recognition in Cortex," *Nature Neuroscience*, vol. 2, p. 7, 1999.
 - [43] H. Han, C. Otto, and A. K. Jain, "Age Estimation from Face Images: Human vs. Machine Performance," in *Proceedings of the International Conference on Biometrics*, 2013, pp. 1-8.
 - [44] T. Kanno, M. Akiba, Y. Teramachi, H. Nagahashi, and T. Agui, "Classification of Age Group Based on Facial Images of Young Males by Using Neural Networks," *IEICE Transaction on Information and System*, vol. 84, pp. 1094-1101, 2001.
-

-
- [45] K. Ueki, T. Hayashida, and T. Kobayashi, "Subspace-Based Age-Group Classification using Facial Images under Various Lighting Conditions," in *Proceedings of the 7th IEEE International Conference on Automatic Face and Gesture Recognition*, 2006, pp. 48-48.
 - [46] S. Yan, H. Wang, X. Tang, and T. S. Huang, "Learning Auto-Structured Regressor from Uncertain Nonnegative Labels," in *Proceedings of the 11th IEEE International Conference on Computer Vision*, 2007, pp. 1-8.
 - [47] S. Yan, H. Wang, J. Liu, X. O. Tang, and T. S. Huang, "Ranking with Uncertain Labels and its Applications," *Frontiers of Computer Science In China*, vol. 1, pp. 407-412, 2007.
 - [48] S. Yan, X. Zhou, M. Liu, M. Hasegawa-Johnson, and T. S. Huang, "Regression from Patch-Kernel," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2008, pp. 1-8.
 - [49] S. K. Zhou, B. Georgescu, X. S. Zhou, and D. Comaniciu, "Image Based Regression Using Boosting Method," in *Preceedings of the 10th IEEE International Conference on Computer Vision*, 2005, pp. 541-548 Vol. 1.
 - [50] B. Ni, Z. Song, and S. Yan, "Web Image Mining Towards Universal Age Estimator," in *Proceedings of the 17th ACM international conference on Multimedia*, Beijing, China, 2009, pp. 85-94.
 - [51] G. Guo, Y. Fu, C. R. Dyer, and T. S. Huang, "Image-Based Human Age Estimation by Manifold Learning and Locally Adjusted Robust Regression," *IEEE Transactions on Image Processing*, vol. 17, pp. 1178-1188, 2008.
 - [52] X. Geng, C. Yin, and Z.-H. Zhou, "Facial Age Estimation by Learning from Label Distributions," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, pp. 2401-2412, 2013.
 - [53] J. G. Daugman, "Uncertainty Relation for Resolution in Space, Spatial Frequency, and Orientation Optimized by Two-dimensional Visual Cortical Filters," *Journal of the Optical Society of America A*, vol. 2, pp. 1160-1169, 1985.
 - [54] P. Kruizinga and N. Petkov, "Nonlinear Operator for Oriented Texture," *IEEE Transactions on Image Processing*, vol. 8, pp. 1395-1407, 1999.
 - [55] Wikipedia. Available: http://en.wikipedia.org/wiki/Gabor_filter
 - [56] J. G. Daugman, "Two-dimensional Spectral Analysis of Cortical Receptive Field Profiles," *Vision Research*, vol. 20, pp. 847-856, 1980.
 - [57] J. P. Jones and L. A. Palmer, "An Evaluation of the Two-dimensional Gabor Filter Model of Simple Receptive Fields in Cat Striate Cortex," *Journal of Neurophysiology*, vol. 57, pp. 1233-1258, 1987.
 - [58] T. Ojala, M. Pietikainen, and D. Harwood, "Performance Evaluation of Texture Measures with Classification based on Kullback Discrimination of Distributions," in *Proceedings of the 12th International Conference on Pattern Recognition*, 1994, pp. 582-585.
-

-
- [59] T. Ojala, "A Comparative Study of Texture Measures with Classification based on Featured Distributions," *Pattern Recognition*, vol. 29, pp. 51-59, 1996.
 - [60] X. Wang, T. X. Han, and S. Yan, "An HOG-LBP Human Detector with Partial Occlusion Handling," in *Proceedings of the 12th IEEE International Conference on Computer Vision*, 2009, pp. 32-39.
 - [61] T. Serre, L. Wolf, and T. Poggio, "Object Recognition with Features Inspired by Visual Cortex," in *the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2005, pp. 994-1000 vol. 2.
 - [62] Y. Freund and R. E. Schapire, "A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting," *Journal of Computer and System Sciences*, vol. 55, pp. 119-139, 1997.
 - [63] H. Schwenk, "Using Boosting to Improve a Hybrid HMM/neural Network Speech Recognizer," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing* 1999, pp. 1009-1012.
 - [64] G.-B. Huang, D. Wang, and Y. Lan, "Extreme Learning Machines: a Survey," *International Journal of Machine Learning and Cybernetics*, vol. 2, pp. 107-122, 2011.
 - [65] G.-B. Huang, L. Chen, and C.-K. Siew, "Universal Approximation using Incremental Constructive Feedforward Networks with Random Hidden Nodes," *IEEE Transactions on Neural Networks*, vol. 17, pp. 879-892, 2006.
 - [66] I. T. Jolliffe, *Principal Component Analysis*: Springer, 2002.
 - [67] H. Abdi and L. J. Williams, "Principal Component Analysis," *Wiley Interdisciplinary Reviews: Computational Statistics*, vol. 2, pp. 433-459, 2010.
 - [68] A. M. Martinez and A. C. Kak, "PCA versus LDA," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, pp. 228-233, 2001.
 - [69] D. Wetcher-Hendricks, *Analyzing Quantitative Data : An Introduction for Social Researchers* Hoboken, N.J. : Wiley, 2011.
 - [70] R. Fisher, "The Use of Multiple Measurements in Taxonomic Problems," *Annals of Eugenics*, vol. 7, pp. 179-188, 1936.
 - [71] M. Belkin and P. Niyogi, "Laplacian Eigenmaps for Dimensionality Reduction and Data Representation," *Neural Computation*, vol. 15, pp. 1373-1396, 2003.
 - [72] S. Vasuhi, "Identification of Human Faces Using Orthogonal Locality Preserving Projections," in *Proceedings of the International Conference on Signal Processing Systems* 2009, pp. 718-722.
 - [73] F. Yun, X. Ye, and T. S. Huang, "Estimating Human Age by Manifold Analysis of Face Pictures and Regression on Aging Features," in *Proceedings of the IEEE International Conference on Multimedia and Expo*, 2007, pp. 1383-1386.
 - [74] M. Y. Eldib and H. M. Onsi, "Web Image Mining Age Estimation Framework," in *Proceedings of ICGST International Journal on Graphics, Vision and Image Processing, GVIP*, 2011, pp. 1-4.
-

-
- [75] F. McCown and M. L. Nelson, "Search Engines and their Public Interfaces: Which APIs are the Most Synchronized?," in *Proceedings of the 16th international conference of World Wide Web* 2007, pp. 1197-1198.
 - [76] J.-G. Wang, E. Sung, and W.-Y. Yau, "Active Learning for Solving the Incomplete Data Problem in Facial Age Classification by the Furthest Nearest-Neighbor Criterion," *IEEE Transactions on Image Processing*, vol. 20, pp. 2049-2062, 2011.
 - [77] G.-B. Huang, X. Ding, and H. Zhou, "Optimization Method Based Extreme Learning Machine for Classification," *Neurocomputing*, vol. 74, pp. 155-163, 2010.