

Année 2014

Num d'ordre: 41465

UNIVERSITÉ LILLE1

THÈSE
pour obtenir le grade de
DOCTEUR,
SPÉCIALITÉ INFORMATIQUE



présentée et soutenue publiquement par

Baiqiang XIA

le 25/11/2014

**LEARNING 3D GEOMETRIC FEATURES FOR
SOFT-BIOMETRICS RECOGNITION**

préparée au sein du laboratoire LIFL

sous la direction de

Mohamed DAOUDI
Boulbaba BEN AMOR

COMPOSITION DU JURY

M. Liming CHEN	Professeur, Ecole Centrale de Lyon	Rapporteur
M. Jean-Luc DUGELAY	Professeur, EURECOM	Rapporteur
M. Luigi LANCIERI	Professeur, Université Lille 1	Examineur
M. Stephano BERRETTI	Associate Professor, Université de Florence (Italie)	Examineur
M. Christophe ROSENBERGER	Professeur, ENSI Caen	Examineur
M. Mohamed DAOUDI	Professeur, Télécom Lille	Directeur
M. Boulbaba BEN AMOR	Maître de conférences, Télécom Lille	Encadrant

PUBLICATIONS

JOURNALS

1. **Baiqiang Xia**, Boulbaba Ben Amor, Hassen Drira, Mohamed Daoudi, and Lahoucine Ballihi. Combining Face Averageness and Symmetry for 3D-based Gender Classification. **Accepted** in *Pattern Recognition*, 2014 (in press).
2. **Baiqiang Xia**, Boulbaba Ben Amor, Mohamed Daoudi. Joint Gender, Ethnicity and Age Estimation from 3D Faces - An Experimental Illustration of their Mutual Correlations. Submitted to *International Journal of Computer Vision (IJCV)*, 2014.

CONFERENCES

1. **Baiqiang Xia**, Boulbaba Ben Amor, Mohamed Daoudi, Hassen Drira. Can 3D Shape of the Face Reveal your Age ? The 9th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISAPP) 2014: 5-13.¹
2. **Baiqiang Xia**, Boulbaba Ben Amor, Di Huang, Mohamed Daoudi, Yunhong Wang, Hassen Drira. Enhancing Gender Classification by Combining 3D and 2D Face Modalities. Proceeding of European Signal Processing Conference (EUSIPCO). 2013: 1-6.
3. **Baiqiang Xia**, Ben Amor Boulbaba, Hassen Drira, Mohamed Daoudi and Lahoucine Ballihi. Gender and 3D Facial Symmetry: What's the Relationship? Proceeding of the IEEE Conference on Automatic Face and Gesture Recognition (FG) 2013. 2013: 1-6.
4. **Baiqiang Xia**, Boulbaba Ben Amor, Mohamed Daoudi. Exploring the magnitude of sexual dimorphism in 3D face based gender classification. ECCV 2014 International Workshop on Soft Biometrics. 2014. (Invited to submit an extension to Pattern Recognition Letters)

¹This paper won the **Best Paper Award** in the area of **Image and Video Understanding** in VISAPP 2014.

ACKNOWLEDGEMENT

I would like to give my appreciations to my thesis director, Pr. Mohamed Daoudi for his professional guidance, support and personal understanding during my PHD study. His guidance and support gave me strong help all the way in the PHD research.

I would like also to thank sincerely my co-advisor Dr. Boulbaba Ben Amor, for his outstanding professional leadership and great personal qualities. His guidance in research enriched me from a beginner of this domain, to a professional researcher. His friendly and encouraging way in guidance and daily communication gave me strong interest and confidence in research.

Special thanks to my PhD committee members for taking the time to participate in the defence, and especially the reviewers of the manuscript for having accepted this significant task: Pr. Liming Chen and Pr. Jean-Luc Dugelay. I also thank the committee Pr. Luigi Lancieri, Dr. Stephano Berretti, and Pr. Christophe Rosenberger. Thank all of them for giving me the honor to present my thesis on this special day despite their highly charged agendas.

A special thank goes to our research buddies Hazem Wannous, Hassen Drira and Jean-Philippe Vandeborre. I would also thank my freinds and colleagues, Rim Slama, Taleb Alashkar, Maxime Devanne, Meng Meng, Vincent Leon, Sebastien Poulmane and Paul Audain Desrosiers for their friendship and support. I would like also to thank the colleagues in the IRIP group of Beihang University for their assistance and collaboration during my short-term exchange.

I also gratefully acknowledge the Chinese government for providing the scholarship during my thesis, and LIFL Laboratory for providing the research essentials and the environment. I thank my friends, people at

TELECOM Lille and LIFL for their helps, services and encouragements. I want to thank all my family for their love, support and encouragements.

Lille, November 05, 2014.

CONTENTS

PUBLICATIONS	v
ACKNOWLEDGEMENT	v
CONTENTS	v
LIST OF FIGURES	viii
1 INTRODUCTION	1
1.1 MOTIVATIONS	3
1.2 THESIS CONTRIBUTIONS	5
1.3 THESIS ORGANIZATION	6
2 RELATED WORK	7
2.1 INTRODUCTION	8
2.2 FACE GENDER RECOGNITION	9
2.2.1 2D based face gender classification	9
2.2.2 3D based face gender classification	12
2.2.3 2D+3D based face gender classification	14
2.3 FACE ETHNICITY CLASSIFICATION	15
2.3.1 2D based face ethnicity classification	15
2.3.2 3D based face ethnicity classification	16
2.3.3 2D+3D based face ethnicity classification	17
2.4 FACE AGE ESTIMATION	17
2.5 JOINT FACIAL SOFT-BIOMETRICS RECOGNITION	22
2.6 CONCLUSIONS	23
3 GEOMETRIC FEATURES EXTRACTION FROM 3D FACE	25
3.1 INTRODUCTION	26

3.2	MORPHOLOGY CUES IN THE FACE	26
3.3	CURVE-BASED SHAPE ANALYSIS FRAMEWORK	28
3.3.1	3D face preprocessing and radial curve extraction	29
3.3.2	Shape analysis of radial curves	30
3.4	GEOMETRIC FEATURE EXTRACTION	33
3.4.1	Feature extraction based on the morphological observations	33
3.4.2	Correlation between features and the facial soft-biometrics	36
3.4.3	Statistical characteristics of the extracted features	39
3.5	CONCLUSIONS	42
4	FACIAL SOFT-BIOMETRICS RECOGNITION	45
4.1	INTRODUCTION	46
4.2	EXPERIMENT SETTINGS	46
4.2.1	Evaluation Dataset and Protocols	46
4.2.2	Random Forest classifier/regressor	47
4.2.3	Dimensionality reduction methods	48
4.2.4	Feature-Level Fusion method	50
4.3	EXPRESSION-DEPENDENT SOFT-BIOMETRIC RECOGNITION	51
4.3.1	Experiments with high dimensional DSF features	51
4.3.2	Experiments with feature dimensionality reduction	55
4.4	EXPRESSION-INDEPENDENT SOFT-BIOMETRIC RECOGNITION	61
4.4.1	PCA-based Expression-Independent experiments	62
4.4.2	CFS-based Expression-Independent Experiments	64
4.4.3	Summary	66
4.5	FEATURE LEVEL FUSION EXPERIMENTS	68
4.5.1	Expression-Dependent Experiments	69
4.5.2	Expression-Independent Experiments	70
4.5.3	Summary	72
4.6	ROBUSTNESS TO FACIAL VARIANTS	72
4.7	ROBUSTNESS TO TRAINING SIZE	75
4.8	COMPARISON WITH THE STATE-OF-THE-ART	77
4.9	CONCLUSION	79
5	JOINT FACIAL SOFT-BIOMETRICS RECOGNITION	81

5.1	INTRODUCTION	82
5.2	CORRELATED FACIAL ATTRIBUTES RECOGNITION	83
5.2.1	Gender Classification Experiments	83
5.2.2	Ethnicity Classification Experiments	85
5.2.3	Age Estimation Experiments	86
5.3	HOW MUCH ARE GENDER, ETHNICITY AND AGE CORRELATED ?	89
5.4	HOW TO BENEFIT FROM THEIR CORRELATION IN REAL-WORLD LIKE APPLICATIONS ?	92
5.5	CONCLUSION	95
6	CONCLUSION	97
	BIBLIOGRAPHY	101

LIST OF FIGURES

3.1	An illustration of the spherical structure of the manifold. . .	31
3.2	3D-Symmetry description from pairwise radial curves. . . .	33
3.3	Illustrations of different DSF features on 3D shape of the face	34
3.4	The average face template	35
3.5	Correlation between facial geometric features and facial soft-biometrics	37
3.6	DSF feature distribution considering different Gender . . .	40
3.7	DSF feature distribution considering different Ethnicity . . .	42
3.8	DSF feature distribution considering different Age	43
4.1	Expression-Dependent Experiments with DSF features . . .	51
4.2	Cumulative Score of age estimation with DSF features . . .	55
4.3	Expression-Dependent Experiments with PCA transformed features	56
4.4	Cumulative Score of age estimation with PCA transformed features	58
4.5	Expression-Dependent Experiments with CSF selected features	59
4.6	Cumulative Score of age estimation with CFS selected features	61
4.7	Expression-Independent Experiments with PCA transformed features	62
4.8	Cumulative Score of age estimation with PCA transformed features	65
4.9	Expression-Independent Experiments with CSF selected features	66
4.10	Cumulative Score of age estimation with CFS selected features	67
4.11	Feature level fusion results for facial soft-biometric recognition	68
4.12	Cululative score of age estimation with fusion of selected features	70
4.13	Gender recognition rate concerning facial Variants	73
4.14	Ethnicity recognition rate concerning facial Variants	74
4.15	Age recognition MAE concerning facial Variants	75
4.16	Recognition accuracy when varying the # of training scans	76
5.1	Gender classification in Expression-Dependent and -Independent settings	84
5.2	Ethnicity classification in Expression-Dependent and -Independent settings	86

5.3	Age classification in Expression-Dependent and -Independent settings	87
5.4	Distribution of the projected LDA features	90
5.5	Recognition results of the three biometrics using different selected features.	91
5.6	Gender classification using automatic recognition results of ethnicity and age.	93
5.7	Ethnicity classification using automatic recognition results of gender and age.	94
5.8	Age estimation using automatic recognition results of gender and ethnicity.	95

INTRODUCTION

1

SOMMAIRE

1.1	MOTIVATIONS	3
1.2	THESIS CONTRIBUTIONS	5
1.3	THESIS ORGANIZATION	6

Soft-biometrics are human characteristics providing physical or behavioral information for people categorization, such as gender, height, weight, age, and ethnicity, beard, skin/eye/hair color, length of arms and legs, etc [56, 7, 18, 19]. Compared to the classical "hard" biometrics, such as fingerprint, face, hand-geometry, iris, retina, palm-print, ear, voice, gait, signature, keystroke dynamics, etc., soft biometrics provide less determinative information which is not necessarily permanent or distinctive for individuals. The soft-biometrics science began with the French criminologist Alphonse Bertillon early in the nineteenth century, who introduced the idea for a personal identification system based on biometric, morphological and anthropometric determinations [83, 18, 19]. He used traits like colors of eye, hair, beard and skin; shape and size of the head; general discriminators like height or weight and also description of indelible marks such as birth marks, scars or tattoos [83, 19]. In [56, 7], *Jain et al.* formally defined the soft-biometric traits as characteristics that provide some information about the individual, but lack in the distinctiveness and permanence to sufficiently differentiate any two individuals. Soft biometrics can be continuous (e.g., age, height and weight, which have continuous values), or discrete (e.g., gender, eye color, ethnicity, etc., which have discrete values). And usually they are easier to capture at distance and do not require cooperation from the subjects [55]. The applications of soft-biometrics include human identification, Human-Machine Interaction, content based image/video retrieval, person re-identification, etc. In [19], *Dantcheva et al.* redefined the soft-biometric traits as physical, behavioral or adhered human characteristics, classifiable in pre-defined human compliant categories which are established and time-proven by humans with the aim of differentiating individuals. In this definition, not only the face and body traits (as in [56, 7]), but also the accessories like glasses and clothes are taken as soft-biometrics.

As a "window to the soul" [32, 103], human faces demonstrate important perceptible cues related to individual soft-biometric traits, such as the gender, age, ethnicity, facial expression, and pose. The common facial soft-biometrics include the age, the gender and the ethnicity, the

skin/eye/hair colors, the existence of beard and mustache [19]. In [57], *Jain et al.* proposed to use facial marks (eg. freckles, moles, and scars) as soft-biometrics to improve face recognition and retrieval performance. Among the facial soft-biometrics, gender, ethnicity and age have attracted more investigations as they don't only convey categorical information of individuals, but also relate closely to anthropometry, demographics and cognitive science. The gender recognition links directly to the study of sexual dimorphism in face. The age estimation relates closely to the aging phenomenon of human beings, which is an important issue and has received a big amount of attention in biological study ¹. These traits are also related to the interpretation of facial attractiveness and beauty.

Human beings acquire the ability of face perception for these traits in early age, and perform the recognition relatively accurately in daily life. Doing face-based gender, ethnicity and age recognition are basic and important tasks in our social interactions. Among the last decades, due to the extensive needs in computer graphics and computer vision fields, automatic face image processing techniques have attracted much attention from both industry and research. Beyond the scope of identity recognition, the recognition of these soft-biometrics have already developed into independent research topics, especially for applications where we don't need to specifically identify the individual. The recognition of gender, ethnicity and age can contribute in many applications, such as in human-computer interaction, security control and surveillance monitoring, content-based indexing and searching, demographic collection and targeted advertising, forensic art, entertainment and cosmetology [32, 77].

1.1 MOTIVATIONS

The present thesis addresses the problem of facial soft-biometrics recognition, specially for gender, ethnicity and age, using the 3D shape of faces. Traditionally, facial soft-biometric recognition researchers used the 2D face images. In 2D images, faces are represented with a 2D light intensity func-

¹The 2009 Nobel Prize in Physiology or Medicine was given to three American scientists who made key discoveries about how living cells age.

tion $f(x,y)$, where x and y are the 2D spatial coordinates and the value of f signifies the brightness (or the color) of the face on this point. The information captured in the 2D images is significantly sensitive to the illumination condition and the head/camera pose changes. Recently, with the advances in 3D imaging techniques, more and more researchers explored the usage of the 3D face scans for soft-biometric recognition. Instead of capturing the intensity information in 2D images, the 3D face scans capture the depth information of faces, which in the end results in a 3D spatial representation of the face. Compared to 2D face images, the 3D faces are able to represent the complete 3D geometry of the face surfaces, and are independent to illumination and robust to head/camera pose changes.

Another motivation for using 3D scans is the relationship between the 3D shape of face and the facial biometrics. Human faces present rich cues in shape for recognizing their Gender, Ethnicity and Age, by their peers. Study in anthropometry [106] has revealed that different gender, ethnicity and age groups have significantly different facial morphology. Specifically, the studies in sexual dimorphism [106, 1, 92] have concluded that male faces usually possess more prominent features than female faces. For Ethnicity, statistics in anthropometry have shown that the morphological differences exist in the craniofacial complex in different ethnicity [26, 30, 75]. The Asian and Non-Asian population convey significantly different facial morphology [106, 69, 4, 6, 26], such as the face width, and the width and height of the nose. In the study of face aging [84, 81], researchers have concluded that, the craniofacial growth is the main change in baby and adolescent face, and the face contour and texture keep changing in the adulthood. Thus, the 3D shape of face encodes rich information of Gender, Ethnicity and Age. The 3D scans contain completely the encoded information in the facial shapes. However, when faces are projected onto a 2D plane in 2D images, the shape information is incomplete and even distorted. Thus, the 2D images can not capture appropriately the geometric cues of Gender, Ethnicity and Age encoded in facial shape.

Thus, considering the merits of of 3D scans than 2D images, and the rich cues in 3D face shape related to Gender, Ethnicity and Age, our work

is established using the 3D face scans. In [51], *Hu et al.* have demonstrated that, with the 3D scans, human observers perform better on both gender recognition and ethnicity recognition tasks than with 2D face images. In biology study [49], researchers have also found that, when considering gender and ethnicity recognition tasks, the usage of 2D face images is limited to full-face view, while the 3D scans are proved to be adaptable to angled views (non-frontal poses).

1.2 THESIS CONTRIBUTIONS

In this thesis, our methodology consists on learning various geometric features from 3D shape of face, for estimating the Gender, Ethnicity and Age from 3D face, and for exploring their correlations. The main contributions can be summarized as following:

- *Expression-Robust 3D facial soft-biometrics recognition using geometric features* – We propose four different and complementary facial descriptors grounding on Shape Analysis of facial radial curves, with which we demonstrate that **the 3D shape of face can reveal our gender, ethnicity and age**. Extensive evaluations on the challenging FRGCv2 dataset demonstrate the effectiveness of the proposed facial attribute recognition approach and its robustness to facial expressions. Our approach is also robust to the size of training data. To our knowledge, the work concerning age estimation is the first work in the literature which studies age estimation using 3D face.
- *Joint facial soft-biometrics recognition* – We propose to explore the usage of the correlations among these soft-biometrics in their recognition tasks. We demonstrate that **gender, ethnicity and age are correlated in the 3D face**, and the correlations are helpful in terms of both the recognition accuracy and the computational cost in each others' recognition tasks. As far as we know, this is the first work in the literature which gives thorough study of the correlations among Gender, Ethnicity and Age from the 3D shape of faces. We discover that the

correlation between ethnicity and age is the strongest among the correlations.

1.3 THESIS ORGANIZATION

The thesis is organized as following. In chapter 2, we lay out the state of the art on face-based Gender, Ethnicity and Age estimation. In chapter 3, we detail our geometric feature extraction strategy and discuss the relationship between the extracted features and the facial soft-biometrics. Chapter 4 presents the experimental evaluation of our facial soft-biometrics recognition method. In chapter 5, we explore the correlation among these soft-biometrics in their recognition tasks. Finally, chapter 6 makes the conclusion of the thesis and explores some perspectives.

RELATED WORK

SOMMAIRE

2.1	INTRODUCTION	8
2.2	FACE GENDER RECOGNITION	9
2.2.1	2D based face gender classification	9
2.2.2	3D based face gender classification	12
2.2.3	2D+3D based face gender classification	14
2.3	FACE ETHNICITY CLASSIFICATION	15
2.3.1	2D based face ethnicity classification	15
2.3.2	3D based face ethnicity classification	16
2.3.3	2D+3D based face ethnicity classification	17
2.4	FACE AGE ESTIMATION	17
2.5	JOINT FACIAL SOFT-BIOMETRICS RECOGNITION	22
2.6	CONCLUSIONS	23

2.1 INTRODUCTION

Gender, Ethnicity, and Age are natural recognizable traits in human faces. In anthropometry [106], it has been revealed that significant facial morphology differences exist in different gender, ethnicity and age groups. When studying the sexual dimorphism [1, 92], researchers have found that male faces usually possess more prominent features than female faces. Male faces usually have more protuberant noses, eyebrows, more prominent chins and jaws. The forehead is more backward sloping, and the distance between top-lip and nose-base is longer. Research presented in [106] has also demonstrated that females are smaller in all the concerned anthropometric measurements. For ethnicity, statistics in anthropometry have shown that morphological differences exist in the craniofacial complex in different ethnicity [26, 30, 75]. In [69, 4], researchers have found that compared with the North America Whites, the Asian population usually have broader faces and noses, farther apart eyes, and exhibit the greatest difference in the anatomical orbital regions (around eyes and eyebrows). In [26], *Farkas et al.* study the head and face of North American Caucasian, African American and Chinese, and identify that the Chinese have the widest face, largest intercanthal width, highest upper lip in relation with the mouth width, and less protruding and wider nose. In the clinical study reported in [6], *Alphonse et al.* have shown that the Caucasians have significantly lower fetal frontomaxillary facial angle (FMFA) measurements than Asians. In [106], 16 anthropometric measurements have been recognized as significantly different between Asian and Caucasian faces. In the study of face aging [84, 81], researchers have concluded that, the craniofacial growth is the main change in baby and adolescent face, which results in the re-sizing and redistribution of facial features. In this period, the bigger is the size of the face, generally the larger is the age. When the craniofacial growth stops at 18-20 years old, the face contour and texture changes become the dominant changes. Young adults tend to have more a triangle shaped face with small amount of wrinkles. In contrast, old adults are usually associated with a U-shaped face with significant wrinkles [84].

In daily life of human beings, we are doing face-based gender, eth-

nicity and age recognition naturally and effectively. In human face analysis using machines, automatic Gender, Ethnicity and Age recognition have become active research areas during the last two decades. Developed solutions could be used in human computer interaction (intelligent user interface, video games, etc.), visual surveillance, collecting demographic statistics for marketing (audience or consumer proportion analysis, etc.), and security industry (access control, etc.). Fundamentally, the proposed approaches comprise of two main stages: a feature extraction stage which makes an appropriate representation of facial images, and a machine learning stage which applies classification/regression algorithms on the extracted feature for attribute recognition. The proposed techniques for recognition of these soft-biometrics differ in (i) the format of data (still 2D images, 2D videos or 3D scans); (ii) the choice of facial representation (features), ranging from simple raw 2D pixels or 3D cloud of points to more complex features, such as Haar-like, LBP and AAM in 2D, and shape index, wavelets and facial curves in 3D; and (iii) design of the classifiers/regressors, for instance the Random Forest, Neural Networks, Support Vector Machine (SVM), and Boosting methods.

2.2 FACE GENDER RECOGNITION

2.2.1 2D based face gender classification

Human faces exhibit clear sexual dimorphism (SD), in terms of masculinity and femininity [58], for recognizing their gender. As stated earlier, researchers have concluded that male faces usually possess more prominent features than female faces [106, 1, 92, 53]. Since the 1990s, automated gender classification has gradually developed as an active research area. Abundant works have been published. The earliest ones are based on 2D images. In [76], *Erno Makinen and Roope Raisamo* compare gender classification performance on frontal 2D intensity face images of different methods (with different input image sizes, normalization methods and classifiers, including Neural Network, Adaboost and SVM). For Neural Network, histogram-equalized image pixels are taken as input. For SVM, they try

face image pixels and LBP features separately as input. And for Adaboost, they use the Haar-like features as input. Experiments are carried out on FERET database and an internet dataset. Results show that statistically no significant difference exists between the performance of classifiers. The database selection, normalization (rotation, scaling, cropping), presence of hair, and the experiment settings account more for the deviation of results. They also propose an arithmetical way to combine the outputs of classifiers with which the classification rate is improved.

Gender recognition results could be influenced by face descriptors. In [101], *Ylioinas et al.* combine Contrast Information (strength of patterns) and Local Binary Patterns (LBP patterns) in VAR/LBP framework for 2D face representation. Experiments are carried out on FRGCv2, FERET and XM2VTS databases with SVM (LIBSVM) using 5-fold cross validation. The best result reported across the three databases is 96.33%, while in contrast with pure LBP representation and SVM it is 92.31%; the result for the same VAR/LBP representation with Adaboost is 81.10%.

Both local and global information are crucial for image perception. In [99], *Yang et al.* propose a Global-Local-Features-fusion (GLFF) approach for gender classification. The outputs of the global-feature descriptor AMM and the local-feature descriptor LBP are fused using sequent selection algorithms. Standard 5-fold cross-validation experiments using SVM are then performed on the FGNet database with best result of 85.02%.

Describable visual attributes are labels given to an image to describe its appearance. In [63], *Kumar et al.* use describable visual attributes for 2D face verification and image search. They extract low-level features according to the labels of face region, pixel data, applied normalization and aggregation. To train each attribute classifier, a discriminating subset of these features is generated through an automatic forward feature selection procedure. Finally, these attribute classifiers are fed together to RBF-SVM to decide the verification and searching results. Gender is treated as one of the visual attributes. On the Columbia Face Database, they achieve a correctness of 91.38% using RBF-SVM on a set of near-frontal faces with

only gender and smiling attributes considered, and 85.8% with 73 attribute classifiers considered.

The robustness against environment variations is an important property for the efficiency of approaches. In [86], *Shan* investigates gender classification on real-world 2D faces. Local Binary Patterns (LBP) are employed as face descriptors and Adaboost is used to select discriminating LBP features. Selected features are then fed to SVM for gender classification. The experiments are carried out on a public database named Labeled Faces in the Wild. The best result reported is 94.81% with 5-fold cross validation, with which the author claims that LBP has good robustness to environment variations. In [14], *Wang et al.* enhance LBP with one of its variants, named Local Circular Patterns (LCP), for gender classification. Other than using the Uniformed LBP patterns, they cluster the LBP codes with the K-means clustering with L_1 distance. They achieve 95.36% classification rate on the FERET database, with linear kernel SVM classifier.

Gender and Ethnicity are naturally correlated facial biometrics. In [33], *Gao et al.* learn ethnicity-specific gender classifiers and achieve higher overall classification results on a collection of 2D images. In [28], *Giovanna et al.* perform ethnicity-specific gender classification with the 2D FERET and TRECVID datasets. **Unexpectedly, they find that the ethnicity information is not helpful in gender classification.**

In limited computational resource contexts, such as the mobiles, the development of resource-limited algorithms is important for applications of computer vision and pattern recognition. In this case, linear classification techniques attract attention due to its simplicity and low computational cost. In [11], *Bekios-Calfa et al.* study 2D-based face gender recognition with Linear Discriminative Analysis (LDA). With a Bayesian classifier assuming Gaussian distribution, experiments have been performed with four different datasets. Results show that linear techniques can achieve similar accuracy to SVM or Boosting classifiers within a large dataset, and the linear approach performs relatively better when training data and computational resources are very scarce.

Different information may contribute to face based gender classifica-

tion differently. In [79], *Perez et al.* use multi-information (gray-scale intensity, edge-map of range image, and LBP texture) for 2D-based frontal face gender classification. With three mutual information based feature selection methods, they perform a set of experiments in three different spatial scales for each type of face information. By fusing the selected features of 6 different experiments, they achieve an accuracy of 99.13% on FERET dataset.

2.2.2 3D based face gender classification

Human faces are approximately bilaterally symmetrical. With this assumption, in [73], *Liu et al.* look into the relationship between facial asymmetry and gender. They impose a 2D grid on 3D face mesh to represent the face with 3D grid points. With the selected symmetry plane which equally separates the face into right and left halves, the distance difference (Euclidean distances to the origin of the cylindrical co-ordinate system) between each point and its corresponding reflected point is calculated as height differences (HD), and the angle difference between their normal vectors is calculated as orientation differences (OD). Thus for each face, a HD-face and a OD-face are generated and presented in matrices. The relationships between gender and overall feature asymmetry are examined by comparing overall mean value in both HD-face and OD-face. They also define a local symmetry measurement named Variance Ratio (VR). With VR, a discriminative low-dimensional subspace of HD- and OD-face feature space are generated and then fed to LDA to investigate the relationship between gender and local feature asymmetry. Results on 111 3D neutral faces of 111 subjects show that statistically significant difference could be observed between genders with overall OD facial asymmetry measurement. With the output of LDA, they achieve 91.16% and 96.22% gender recognition rate in testing on HD-face and on OD-face respectively. Results also support early claims in psychology research that statistically male faces possess larger amount of asymmetry than female [69].

Statistically there are differences between geometry facial features of different gender, such as in the hairline, forehead, eyebrows, eyes, cheeks,

nose, mouth, chin, jaw, neck, skin, beard regions [1]. In [47], *Han et al.* present a geometry feature based approach for 3D-face gender classification. Volume and area of forehead, and their corresponding ratio to nose, eyebrows, cheeks and lips are defined to generate feature vectors. RBF-SVM is then applied to these feature vectors to classify gender. 61 frontal 3D face meshes are selected from GavabDB database for experiment. They perform 5 experiments where each experiment contains 48 faces for training and 13 for testing. The average classification rate reported is 82.56%.

In [50], *Hu et al.* propose a fusion-based gender classification method for 3D frontal faces. Each 3D face shape is separated into four face regions using face landmarks. With the extracted features from each region, the classifications are done using SVM on a subset of the UND dataset and another in-house dataset. Results show that the upper region of the face contains the highest amount of discriminative gender information. Fusion is applied to the results of four face regions and the best result reported is 94.3%. Their experiments only involve neutral faces. No attention is given to facial expressions.

In [9], *Ballihi et al.* extract facial curves (26 iso-level curves and 40 radial curves) from 3D faces for gender classification. The features are extracted from lengths of geodesics between facial curves from a given face to the Male and Female templates computed using the Karcher Mean Algorithm. The Adaboost algorithm is then used to select salient facial curves. They obtain a classification rate of 84.12% with the Nearest Neighbor classifier when using the 466 earliest scans of the FRGCv2 dataset as the testing set. They also perform a standard 10-fold cross-validation for the 466 earliest scans of FRGCv2, and obtain 86.05% with Adaboost.

In [91], *Toderici et al.* employ MDS (Multi-Dimensional Scaling) and wavelets on 3D face meshes for gender classification. They select 1121 scans of Asian subjects and 2554 scans of White subjects from FRGCv2 for ethnicity and gender classification. Experiments are carried out subject-independently with no common subject used in the testing stage of 10-fold cross validation. With polynomial kernel SVM, they achieve 93% gender classification rate with the unsupervised MDS approach, and 94%

classification rate with the wavelets-based approach. Both approaches significantly outperform the kNN and kernel-kNN approaches. In their experiment, the authors consider only the *Asian* and *White* ethnicity classes and leave out 332 scans of 48 subjects of other ethnicity groups in FRGCv2 dataset.

In [36], *Gilani et al.* automatically detect the biologically significant facial landmarks and calculate the euclidean and geodesic distances between them as face features. Minimal-Redundancy-Maximal-Relevance (MRMR) algorithm based feature selection is used to find salient features. In 10-fold cross-validation with a LDA classifier, they achieve 96.12% gender classification rate with the combination of euclidean and geodesic features. Taking individually, the geodesic features outperform the euclidean features. It indicates that the geodesic features capture better the shape information of 3D faces. Their approach requires accurate detection of a set of facial landmarks.

2.2.3 2D+3D based face gender classification

Range and intensity modalities of face provide different cues of demographic information. In [74], *Lu et al.* provide an integration scheme for range and intensity modalities to classify ethnicity (Asian and Non-Asian) and gender (Male and Female). SVM is used to extract posterior probabilities from normalized range and intensity images. Posterior probability values of range and intensity are then fused with equal weight and compared directly to classify ethnicity and gender. A mixture of two frontal 3D face databases (UND and MSU databases) is used in their experiments. The best gender classification result using 10-fold cross-validation reported is 91%.

In [95], *Wu et al.* use the 2.5D facial surface normals (needle-maps) recovered with Shape From Shading (SFS) from intensity images for gender classification. The recovered needle-maps presented in PGA (Principle Geodesic Analysis) parameters not only contain facial shape information, but also the image intensity implicitly. Training feature vectors are extracted by LDA from these needle-maps and then used in constructing

Gaussian models to tell the gender of a query face. 260 2D frontal face images are selected from the UND Database. Experiments are done 10 times with 200 faces randomly selected for training and the left 60 faces for testing. The best average gender recognition rate reported is 93.6% with both shape and texture accounted.

In [54], *Huynh et al.* fuse the Gradient-LBP features from range images, and the Uniform LBP features on 2D gray images, for gender classification. With RBF kernel SVM, they achieve 90.38% gender classification rate on the EURECOM Kinect Face Dataset (concerning three scenarios, Neutral, Smile and Light On), and 96.70% on the Texas 3DFR dataset. For the experiments on both the datasets, the first half of male and female subjects are used for training, and the second half are used for testing.

Recently, in [52], *Huang et al.* fuse the decisions of Adaboost classifier from both texture and range images for gender classification, using the Local Circular Patterns (LCP) features. With 3676 face samples of the 99 Asian and 319 White subjects of the FRGCv2 dataset, they achieve 95.50% correct gender classification rate in 10-fold subject-independent cross-validation. Results show the advantage of combining both the texture and depth information in face based gender classification, than with only the texture or the depth information.

2.3 FACE ETHNICITY CLASSIFICATION

2.3.1 2D based face ethnicity classification

Unlike the identity or gender, the ethnic categories are loosely defined classes [75]. In [3], three diversities of ethnicity are recognized, namely Caucasian, Mongolian and Negroid. In [15], *Coon* classifies ethnicity into four major groups, namely white/Caucasian, Mongoloid/Asian, Negroid/Black, and Australoid. The national statistics office of the United Kingdom recommends five options of ethnicity groups for civilians, including White, Asian, Black, Mixed and Others¹. In automated face studies, ethnicity is usually interpreted into 2-4 classes [21, 44, 100, 75], and

¹<http://www.ons.gov.uk/ons/guide-method/measuring-equality/equality/ethnic-nat-identity-religion/ethnic-group/>

very limited exploration has been done for ethnicity classification, than for gender classification. In [44], *Gutta et al.* propose the first work, with a hybrid architecture of the Ensemble Radial Basis Function (ERBF) networks and the decision trees. They consider four ethnicity groups, namely Caucasian, Asian, Oriental and African, and achieve 94% ethnicity classification rate. In [21], *Demirkus et al.* classify ethnicity with SVM classifier in 4-fold cross validation on pixel intensity and Biologically Inspired Model (BIM) features. Considering three ethnicity groups, namely African American, Caucasian and Asian, with the pixel intensity-based features, they achieve 92.6% ethnicity classification rate on 600 face images. With the BIM features, they achieve 85.0% classification rate on 200 images, considering Asian and Non-Asian groups. In [100], *Yang et al.* perform ethnicity classification with LBP features, considering the Asian and Non-Asian groups. With images of a snapshot dataset and FERET dataset in training, they achieve 96.99% classification rate with AdaBoost classifier when testing on the PIE dataset. In [75], *Lu et al.* classify 2D images into Asian and Non-Asian groups with Multi-Scale LDA classifiers. In [28], *Giovanna et al.* perform gender-specific ethnicity classification with the 2D FERET and TRECVID datasets. **Again unexpectedly, they find that the gender information is not helpful in ethnicity classification.**

2.3.2 3D based face ethnicity classification

For the shape-based works, in [104], *Zhong et al.* perform fuzzy ethnicity recognition of eastern and western groups on 3D FRGCv2 dataset, with the Learned Visual Codebook (LVC) derived from histograms of Gabor features. Experimental results show that the facial expressions give strong influence to the ethnicity classification performance. In [91], *Toderici et al.* extract features using the wavelets and the MDS on the Asian and White subsets of the 3D FRGCv2 dataset. With Polynomial kernel SVM in 10-fold cross-validation, they achieve 99.5% classification rate with the MDS method, and 97.5% with the wavelet method.

2.3.3 2D+3D based face ethnicity classification

Concerning the texture and shape-based works, in [74], *Lu et al.* use the ensemble of 2D and 3D scans from the UND and MSU datasets, considering Asian and Non-Asian groups. They achieve 98% classification rate using 10-fold cross-validation, which outperforms the results using only the 2D intensity or 3D range images. In [22], *Ding, et al.* learns a compact set of features from the Oriented Gradient Maps (OGMs) which capture the local geometry and texture variations of entire faces. Experimenting on the FRGCv2 dataset, they reached 98.3% classification rate to distinguish Asians from non-Asians, with 80% samples used in the training set. In [52], *Huang et al.* fuse the decisions of Adaboost from both texture and range images for ethnicity classification, using the Local Circular Patterns (LCP) features. With 3676 face samples of the 99 Asian and 319 White subjects of the FRGCv2 dataset, they achieve 99.60% correct gender classification rate in 10-fold subject-independent cross-validation. Results show the advantage of combining both the texture and depth information in face based gender classification, than with only the texture or the depth information.

2.4 FACE AGE ESTIMATION

Face age estimation performs important social roles in human-to-human communication. Studies in cognitive psychology, presented as a review in [84], have discovered that human beings develop the ability of face age estimation naturally in early life, and can be fairly accurate in deciding the age or age group with a given face. These studies, based on subjective age estimation given to face images from human participants, have also found that multiple cues contribute to age estimation, including the holistic face features (like the outline of the face, face shape and texture, etc.), local face features (like the eyes, nose, the forehead, etc.) and their configuration (like the bilateral symmetry of the face [16]). Whereas, claims have also been given that individuals are not sufficiently reliable to make

fine-grained age distinctions, and individuals age estimation suffers from the subjective individual factors and contextual social factors.

The aging process is a cumulative, uncontrollable and personalized slow process, influenced by intrinsic factors like the gene and gender, and extrinsic factors like lifestyle, expressions, environment and sociality [32, 46]. The appearance and anatomy of human faces changes remarkably with the progress of aging [68]. The general pattern of the aging process differs in faces of different person (personalized or identity-specific), in faces of different age (age-specific), in faces of different gender (gender-specific), and in different facial components [32, 84, 42, 78, 39]. Typically, the craniofacial growth (bone movement and growth) takes place during childhood, and stops around the age of 20, which leads to the re-sizing and re-distribution of facial regions, such as the forehead, eyes, nose, cheeks, lips, and the chin. From adulthood to old age, face changes mainly in the skin, such as the color changes (usually darker and with more color changes) and the texture changes (appearance of wrinkles). The shape changes of faces continues from adulthood to old age. With the droops and sags of facial muscle and skin, the faces are tend to be more a shape of trapezoid or rectangle in old faces, while the typical adult faces are more of a U-shaped or upside-down-triangle [84].

Automatic face age estimation is to label a face image with the exact age or age group objectively by machine. With the rapid advances in computer vision and machine learning, recently, automatic face age estimation have become particularly prevalent because of its explosive emerging and promising real-world applications, such as electronic customer relationship management, age-specific human-computer-interaction, age-specific access control and surveillance, law enforcement (e.g., detecting child-pornography, forensic), biometrics (e.g., age-invariant person identification [78]), entertainment (e.g., cartoon film production, automatic album management), and cosmetology. Compared to human age estimation, automatic age estimation yields better performance as demonstrated in [46].

The performance of age estimation is typically measured by the mean absolute error (MAE) and the cumulative score (CS). The MAE is defined

as the average of the absolute errors between the estimated age and the ground truth age, while the CS, proposed firstly by [34] in age estimation, shows the percentage of cases among the test set where the absolute age estimation error is less than a threshold. The CS measure is regarded as a more representative measure in relation with the performance of an age estimator [66].

As pointed in [84, 81], the earliest age estimation works used the mathematical cardioidal strain model, derived from face anthropometry that measures directly the sizes and proportions in human faces, to describe the craniofacial growth. These approaches are useful for young ages, but not appropriate for adults. After this, abundant works exploiting 2D images have been published in the literature with more elaborated approaches. Different with the comprehensive surveys given by [84, 81], which categorize the literature concerning different aging modeling techniques, we represent the literature with the different ideas underlying these technical solutions. Based on the previous statements, we describe the face appearance as a function of multiple factors, including the age, the intrinsic factors (permanent factors like gene, gender, ethnicity, identity, etc.), and the extrinsic factors (temporary factors like lifestyle, health, sociality, expression, pose,

A. General aging patterns in face appearance. Essentially, face age estimation is to estimate the age of a subject by the aging patterns shown visually in the appearance. To analyze the appearance given in the face image is the basic ways to estimate the age. In the literature of age estimation, works were carried out with several different perceptions of the general aging patterns in face appearance. As aging exhibits similar patterns among different person, several approaches have been designed to learn the general public-level aging patterns in face appearance for age estimation. The most representative ones are the Active-Appearance-Model (AAM) based approaches, the manifold embedding approaches, and the Biologically-Inspired-Feature (BIF) based approaches. The common idea underlying these approaches is to project a face (linearly or non-linearly) into a subspace, to have a low dimensional representation. Respectively,

(i) [68, 67] use an Active Appearance Model (AAM) based scheme for projecting face images linearly into a low dimensional space. The AAM was initially proposed by [17], in which each face is represented by its shape and texture deviations to the mean face with a set of model parameters. Age estimation results with a quadratic regressor show that the generic aging patterns work well for age estimation. Moreover, [67] illustrates that different face parameters obtained from training are responsible for different changes in lighting, pose, expression, and individual appearance. Considering that these parameters work well for age estimation, we can conclude that these face co-variants are influential in age estimation.

(ii) The goal of manifold embedding approaches is to embed the original high dimensional face data in a lower-dimensional subspace by linear or non-linear projection, and take the embedding parameters as face representation. In the work of [39, 38], the authors extract age related features from 2D images with a linear manifold embedding method, named Orthogonal Locality Preserving Projections (OLPP). [70] learns age manifold with both local preserving requirements and ordinal requirements to enhance age estimation performance [96] projects each face as a point on the Grassmann Manifold with the standard SVD method, then the tangent vector on these points of the manifold are taken as features for age estimation.

(iii) Inspired by a feed-forward path theory in cortex for visual processing, [42] introduces the biologically inspired features (BIF) for face age estimation. After filtering an image with a Gabor filter and a standard deviation based filter consecutively, the obtained features are processed with PCA to generate lower-dimension BIF features. The results demonstrate the effectiveness and robustness of bio-inspired features in encoding the generic aging patterns. Beyond the public-level aging patterns, there could be some less generic aging patterns when dealing with a subset of faces, such as a group of faces with high similarity, or a temporal sequence of face images for the same person. Based on the observation that similar faces tend to age similarly, [67, 68] present an appearance-specific strategy for age estimation. Faces are firstly clustered into groups considering their inter similarity, then training is performed on each group separately

to learn a set of appearance-specific age estimators. Given a previously unseen face, the first step is to assign it to the most appropriate group, then the corresponding age estimator makes the age estimation. Experimental results show that the group-level aging patterns are more accurate in age estimation compared with the generic-aging patterns. In case there is no similar enough face image for a testing face image in the database, [68] presents a weighted-appearance-specific which also yield fine performance. As different individual ages differently, [35, 34] propose the Aging-Pattern-Subspace (AGES), which studies the individual-level aging patterns from a temporal sequence of images of an individual ordered by time. For a test face, the aging pattern and the age is determined by the projection in the subspace that has the least reconstruction error. Experiments confirm that individual aging patterns contributes to age estimation. As facial components age differently, the component-level aging patterns are studied for age estimation. [90] represents faces with a hierarchical And-Or Graph. Face aging is then modeled as a Markov process on the graphs and the learned parameters of the model are used for age estimation. They find that the forehead and eye regions are the most informative for age estimation, which is also supported by the conclusion of [46] using the BIF features.

B. Considering the intrinsic/extrinsic factors in facial aging. As stated at the beginning of this introduction, the appearance of face is influenced by intrinsic factors like the gene, gender, and extrinsic factors like lifestyle, expressions, environment and sociality [32, 46]. Several studies have given consideration of the influences of these factors in age estimation with enhanced age estimation performance reported. Specifically, thinking that faces age differently in different age, age-specific approaches are adopted by [67], where age estimation is obtained by using a global age classifier first, then adjusting the estimated age by a local classifier which operates within a specific age range. Similarly, [39, 38] propose a Locally Adjusted Robust Regressor (LARR) for age estimation, which begins with a SVR-based global age regression, then followed by a local SVM-based classification that adjusts the age estimation in a local age range. All of these

age-specific approaches have achieved better performance compared with their corresponding approaches without local adjustment. Considering that different gender ages differently with age, [61, 81, 38, 64] carry out age estimation on male and female groups separately. Considering the individual lifestyle, [68] encode this information together with facial appearance in age estimation, and demonstrated that the importance of lifestyle in determining the most appropriate aging function of a new individual. [61] gives weights to different lighting conditions for illumination-robust face age estimation. [70] gives consideration of the feature redundancy and used feature selection to enhance age estimation. In addition, [81] gives consideration of the feature redundancy and uses feature selection to enhance age estimation. All of the previous works considering the intrinsic/extrinsic factors have gained better age estimation performance in comparison of using face appearance only.

2.5 JOINT FACIAL SOFT-BIOMETRICS RECOGNITION

Besides the existence of these soft-biometrics in the face, Gender, Ethnicity and Age also interact with each other in characterizing the face shape [106]. For example, according to the anthropometric studies above, the shape of the nose is influenced by all the three soft-biometrics. In public perception, female faces usually look smoother and younger than male faces [2], and the Asian faces usually look younger than Non-Asian faces [102]. In [37], *Vignali et al.* have demonstrated both visually and quantitatively that ethnicity and gender are correlated to some extent in 3D face. In [33], *Gao et al.* have found that the gender classifier trained on a specific ethnicity could not get good generalization ability on other ethnicity. In the study of human perception in [37], when the gender information is subtracted from the faces, the human ethnicity classification performance is recognizably lower.

In the literature of facial soft-biometric recognition, several works have been done regarding the correlations among these soft-biometrics. As stated previously, some 2D texture-based works perform ethnicity-specific gender classification [33, 28], gender-specific ethnicity classification [28],

gender-specific age estimation [81, 38, 64, 61, 43], and ethnicity-specific age estimation [43]. Only [43] reports results on ethnicity&gender-specific age estimation. To the best of our knowledge, the age-specific gender recognition, age-specific ethnicity recognition, gender&age-specific ethnicity recognition, and ethnicity&age-specific gender recognition have never been addressed in the literature. In [41, 40], *Guo et al.* perform age, ethnicity (black and white) and gender estimation jointly with the Canonical Correlation Analysis (CCA) method and the Partial Least Square Regression (PLSR) method, using a multi-label regression strategy. These works show promising ways of investigating age, gender and ethnicity information together. These works don't figure out exactly how influential are the correlations of these soft-biometrics in each others' recognition tasks.

2.6 CONCLUSIONS

In this chapter, we have outlined the related works for face-based Gender, Ethnicity and Age recognition. The majority of existing state-of-the-art works use the 2D face images, especially for Ethnicity and Age recognition. Although the 2D-based approaches have demonstrated their effectiveness in the related recognition tasks, to some extent, they are still facing the challenges of illumination condition and head/camera pose changes. Compared to 2D images, the 3D scans have shown their robustness to these challenges, and the advantages in capturing the complete 3D shape information of the faces. Thus, we are motivated to use the 3D face scans in facial soft-biometrics recognition. Moreover, in the state-of-the-art, despite the fact that Gender, Ethnicity and Age co-exist and correlated naturally in face, only a few 2D-based works have given consideration of the correlations among these soft-biometrics in the recognition tasks. No work has been reported with thorough examination of their correlations in the recognition tasks. Although in [41, 40], *Guo et al.* perform age, ethnicity (black and white) and gender estimation jointly, these works have used a multi-label regression strategy which aims at recognizing these soft-biometrics together. Thus, these works don't figure out exactly how influential are the correlations of these soft-biometrics in

each others' recognition tasks. In the next chapters, we will present our approaches for tackling these tasks, which includes the methodology for geometric feature extraction, and the experimental evaluation and analysis with various machine learning techniques.

GEOMETRIC FEATURES EXTRACTION FROM 3D FACE

SOMMAIRE

3.1	INTRODUCTION	26
3.2	MORPHOLOGY CUES IN THE FACE	26
3.3	CURVE-BASED SHAPE ANALYSIS FRAMEWORK	28
3.3.1	3D face preprocessing and radial curve extraction	29
3.3.2	Shape analysis of radial curves	30
3.4	GEOMETRIC FEATURE EXTRACTION	33
3.4.1	Feature extraction based on the morphological observations	33
3.4.2	Correlation between features and the facial soft-biometrics	36
3.4.3	Statistical characteristics of the extracted features	39
3.5	CONCLUSIONS	42

3.1 INTRODUCTION

Feature extraction consists in transforming arbitrary data, such as text, images, videos, into numerical features usable for machine learning tasks, such as detection, retrieval, recognition, classification, regression or clustering. Our goal in this chapter is to extract informative geometric features from 3D faces, which have the capability of characterizing the concerning facial soft-biometrics. To this end, we first introduce a set of facial morphology cues, namely the Averageness of the face, the bilateral Symmetry of the face, the Gradient information (local shape changes) of the face and the Spatial information (global shape changes) of the face. We will show that these morphological cues relate closely to Gender, Ethnicity and Age in the face. Then we propose a mathematical representation that densely captures these morphology cues, which results into four types of face descriptions (features) for recognition of the concerning facial soft-biometrics.

3.2 MORPHOLOGY CUES IN THE FACE

In this section, we propose to extract geometric features in consideration of four types of high-level facial morphology cues, namely the Averageness of the face, the bilateral Symmetry of the face, the Gradient information of the face and the Spatial information of the face. These morphology perspectives are closely related to human perception of face shape, and the soft-biometrics conveyed in the face shape.

The **Averageness** of facial surface has strong relation to Gender, Ethnicity and Age. In face perception, researchers have revealed that facial sexual dimorphism relates closely with anthropometric cues, such as the facial distinctiveness (the converse to averageness) [10], and the bilateral asymmetry [72]. In sexual selection, sexual dimorphism, face averageness and facial symmetry serve as covariants in judging the perceived health of potential mates [87, 82, 71], and also the attractiveness of face [62, 59]. As stated earlier, concerning face averageness, the male faces usually possess more prominent features than female faces [106, 1, 92, 53]. Statistical

studies on head and face of American and Chinese adults reported in [60, 105, 25] have confirmed this point with various face and head measurements. For Ethnicity, the Asian and Non-Asian population convey significantly different facial morphology [106, 69, 4, 6, 26], such as the face width, and the width and height of the nose. The Statistical studies in [60, 105, 25] have also confirmed the facial morphological differences between Asian and Non-Asian population. For Age, in [20, 5], researchers find that exaggerating the distinctiveness in a face produce an increase in the apparent age of the face.

Face **Symmetry** is another high-level geometric cue for indicating Gender, Ethnicity and Age. For Gender, in [72], *Little et al.* reveal that the symmetry and sexual dimorphism from faces are related in humans, and suggest that they are biologically linked during face development. In [89], *Steven et al.* find that the masculinization of the face significantly covaries with the fluctuating asymmetry in men's face and body. For Ethnicity, in [72], *Little et al.* also demonstrate that the Hadza and Europeans populations have different amount of facial asymmetry. For Age, in [16], *Clinton et al.* find that increasing age is associated with a higher degree of facial asymmetry in 3D face surfaces. They measure the root mean square deviation (RMSD) between native and reflected surfaces, and find that 18% of the symmetry variation is accounted for by age, and that asymmetry increase by 4% each decade.

In addition to facial averageness and symmetry, the global changes in facial shape denoted by **Spatiality** of the face relate closely with sexual dimorphism in a face. As demonstrated in [60, 105], sexual dimorphism exhibits inequally in magnitude in different spatial parts of the face. In [8], *Ashok et al.* find that the face features contribute much more than the nose and head features towards sexual dimorphism in the face. For ethnicity, Asian and Non-Asian faces have different morphological differences in different facial parts (Table 7 in [105]). For Age, it's well-known that different facial parts age differently [90, 65]. The internal part of the face and in particular the areas around the eyes are recognized as the most

significant for automatic age estimation [65]. Thus, the spatiality of face shape has close relationship with the concerning facial soft-biometrics.

Another important morphological cue in face shape is the local changes in the shape, termed **Gradient**. For Gender, the interpretation of gender relates closely to the shape gradient which signifies the local shape consistency. It has been revealed that sexual dimorphism demonstrates the developmental stability [71] in face. For Ethnicity, facial features in the Asian and Non-Asian groups have different magnitude of prominence. The Asians usually have wider face and nose, and less protruding nose, than the Non-Asian faces. For Age, the shape gradient captures the wrinkles which are very important in face aging perception [71, 48]

Thus, considering the facial soft-biometrics are closely related to these morphology cues of face, we propose four descriptions for Gender, Ethnicity and Age recognition, which reflect the Averageness ($3D-avg.$), the bilateral Symmetry ($3D-sym.$), the local Gradient ($3D-grad.$), and the global Spatiality ($3D-spat.$) of the 3D face. These features grounding on Riemannian shape analysis of 3D facial curves are able to capture densely the shape deformation on each point. They were first proposed in [12, 23] for facial expression recognition from dynamic flows of 3D faces, called 4D faces.

3.3 CURVE-BASED SHAPE ANALYSIS FRAMEWORK

In this section, we describe four different and complimentary geometric descriptions extracted from the 3D shape of the face. These descriptions are densely extracted using shape analysis of 3D radial curves of the face. Earlier studies on 3D face recognition [24] and 4D expression recognition [12, 23] have been successfully conducted. Although it's common shape analysis framework with [24, 12, 23], our work presents many methodological and application contributions compared to them, which will be highlighted later in this section.

3.3.1 3D face preprocessing and radial curve extraction

Since there are holes, hair and spikes in the raw 3D face scans, preprocessing is necessary to limit their influences. Following the studies in [24, 12, 23], the 3D scans are preprocessed with hole filling, facial part cropping and 3D mesh smoothing, together with nose tip detection and pose normalization. Firstly, through boundary detection, link-up and triangulation, the holes are filled in each mesh. Secondly, since we are working on frontal faces, the nose tip for each mesh is detected with a simple algorithm which locates the nearest point to the 3D scanner. Then the mesh is cropped with a sphere centered at nose tip to discard the hair region. Finally a smoothing filter is used to distribute evenly the 3D vertices. To allow appropriate shape comparison of 3D faces, all the scans are aligned with the same 3D scan using the Iterative Closest Point (ICP) algorithm, to normalize their poses. In practice, as we work on the 3D FRGCv2 dataset, we align all the scans with the first 3D scan in the dataset. With the pipeline described above, a set of 4005 scans over the 4007 scans of FRGCv2 dataset [80] are successively preprocessed.

After preprocessing, following [24], a collection of radial open curves stemming from the nose tip are defined over each face to give an accurate approximation. The radial curve that makes an clockwise angle $\alpha \in [0, 2\pi]$ with the radial curve which passes through the forehead is denoted as β_α , and the neighbor curve of β_α that has an angle increase of $\Delta\alpha$ is denoted as $\beta_{\alpha+\Delta\alpha}$. With equally interpolated α in the range of $[0, 2\pi]$, it results in $S \approx \cup_\alpha \beta_\alpha$, where S and $\beta_{\alpha, \alpha \in [0, 2\pi]}$ denote the preprocessed facial surface and the collection of radial curves, respectively. To allow appropriate shape analysis of these curves, we adopt the mathematical representation proposed in [88], termed the *Square Root Velocity* (SRV) function. This representation allows the registration and analysis of elastic curves under the simple \mathbb{L}^2 metric. In the next, we briefly illustrate the essentials of the mathematical framework for shape analysis of continuous parameterized open curves, because the present thesis builds on these ideas.

3.3.2 Shape analysis of radial curves

Our goal here is to effectively quantify the pairwise shapes differences between facial curves, which requires an accurate curve registration step first to match similar morphological parts in the curves, and then quantify densely their shape difference. As we know, the facial curves are non-rigid and possess elastic properties, such as stretching and bending. Other variabilities such as rotation, translation and re-parameterization also influence the registration and the comparison of the curves. To deal with these challenges, in the present thesis, we adopt a Riemannian shape analysis framework, which has the advantage of performing elastic curve registration and comparison leaded jointly under the \mathbb{L}^2 metric (\mathbb{L}^2 inner product).

An accurate registration step of facial curves is required before quantifying their divergence in shapes. For example, to densely extract the bilateral symmetry of a given face, it is important, for pairwise right-left curves, to match each other such that the peaks match with the peaks and valleys go with the valleys in the pair of curves. To this end, we start by considering a given facial curve as a continuous parameterized function $\beta(t) \in \mathbb{R}^3$, $t \in [0, 1]$. β is first represented by its *Square-Root Velocity Function* SRV, q , according to Eq. (3.1):

$$q(t) = \frac{\dot{\beta}(t)}{\sqrt{|\dot{\beta}(t)|}}, t \in [0, 1], \quad (3.1)$$

where $|\cdot|$ denotes the \mathbb{R}^3 Euclidean norm. Since q enrolls the derivative of the curve, it is invariant to the translation variability of the curve. Then with the \mathbb{L}^2 -norm of q scaled to 1 ($\|q\| = 1$), the scaling variability of the curve is handled. Furtherly, with the SRV representation, it has been proved in [88] that, the elastic metric to compare shapes of curves is invariant to rotation and re-parameterization variabilities, corresponds to the standard \mathbb{L}^2 metric. With re-parameterization, for two curves β_1 and β_2 (represented as q_1 and q_2), the quantities $\|q_1 - q_2\|$ and $\|\sqrt{\gamma}(q_1 \circ \gamma) - \sqrt{\gamma}(q_2 \circ \gamma)\|$ are the same under the \mathbb{L}^2 -norm. Here q de-

notes the SRV function representation of a curve β , and $\sqrt{\gamma}(q \circ \gamma)$ denotes the SRV function of the same curve but with a new parameterization of β dictated by a non-linear orientation-preserving re-parameterization function $\gamma, \gamma : [0, 1] \rightarrow [0, 1]$. This latter property is important as it makes the registration of the curves easier. In fact, it allows to consider one of the curves to be registered as reference and search for the γ^* which optimally registers them with the $\operatorname{argmin}_{\gamma \in \Gamma} (\|q_1 - \sqrt{\gamma}(q_2 \circ \gamma)\|)$. This optimization step allows curves registration by re-parameterization and it is resolved by Dynamic Programming, as described in [88] for general \mathbb{R}^n curves. Thus, the resulted curve registration method allows to establish accurately the morphological correspondence between the curves.

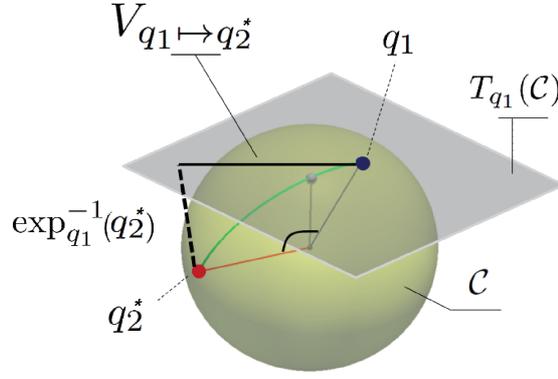


Figure 3.1 – An illustration of the spherical structure of the manifold \mathcal{C} , q_1 (blue) and q_2^* (red) are elements of \mathcal{C} . The geodesic connecting them (green path), $T_{q_1}(\mathcal{C})$ (gray) is the tangent space of \mathcal{C} on q_1 , and the shooting vector $V_{q_1 \rightarrow q_2^*}$ is obtained by $\exp_{q_1}^{-1}(q_2^*)$.

After curve registration, q_2 is re-parameterized optimally to q_2^* , which provides the best registration with q_1 . The following step is to give a measure to the shape differences between curves. In this work, we are interested in a quantity, derived from this mathematical framework, which is capable to capture the pairwise shape differences of radial curves. Formally, as the \mathbb{L}^2 -norm of $q(t)$ has been scaled to 1 ($\|q\| = 1$), the space of such functions: $\mathcal{C} = \{q : [0, 1] \rightarrow \mathbb{R}^3, \|q\| = 1\}$ becomes a Riemannian manifold and has a spherical structure in the Hilbert space $\mathbb{L}^2([0, 1], \mathbb{R}^3)$. According to the spherical structure of \mathcal{C} illustrated in Fig. 3.1, the geodesic path connecting any two points (SVRs of two curves) on the manifold is simply given by the minor arc of the great circle on \mathcal{C} that con-

nects them. Given two curves β_1 and β_2 represented as q_1 and q_2^* on the manifold after registration, the geodesic path $\psi^* : [0, 1] \rightarrow \mathcal{C}$ connecting q_1, q_2^* is given analytically by Eq. (3.2):

$$\psi^*(\tau) = \frac{1}{\sin(\theta)} (\sin((1-\tau)\theta)q_1 + \sin(\theta\tau)q_2^*), \sin(\theta) \neq 0, \quad (3.2)$$

where $\theta = d_{\mathcal{C}}(q_1, q_2^*) = \cos^{-1}(\langle q_1, q_2^* \rangle)$ denotes the distance between q_1 and q_2^* . If $\sin(\theta) = 0$, it means the distance is null, in other words $q_1 = q_2^*$. In this case, $\psi^* : [0, 1] \rightarrow \mathcal{C}$ is given by Eq. (3.3):

$$\psi^*(\tau) = q_1 = q_2^*, \quad (3.3)$$

Knowing that along the geodesic path, the co-variant derivative of its tangent vector field is always equal to 0. The tangent vectors in this field can be obtained simply by parallel transport of the initial tangent vector along the geodesic. Thus, the initial tangent vector, denoted as $V_{q_1 \rightarrow q_2^*}$, is sufficient to represent this vector field and the geodesic path. In our work, we propose to use the initial shooting vector $V_{q_1 \rightarrow q_2^*}$, element of the tangent space on q_1 to the manifold \mathcal{C} , $T_{q_1}(\mathcal{C})$, to capture the shape difference between q_1 and q_2^* . Here again, due to the spherical structure of \mathcal{C} , $V_{q_1 \rightarrow q_2^*}$ is easy to compute using the inverse exponential map given by:

$$\begin{aligned} V_{q_1 \rightarrow q_2} &= \exp_{q_1}^{-1}(q_2^*) \\ &= \frac{\theta}{\sin(\theta)}(q_2^* - \cos(\theta)q_1) \end{aligned} \quad (3.4)$$

where $\theta = \cos^{-1}(\langle q_1, q_2^* \rangle)$ is the length of the geodesic path connecting q_1 to q_2^* , and $q_2^* = (q_2 \circ \gamma^*)\sqrt{\dot{\gamma}^*}$ is the optimal registration of q_2 using the optimal re-parameterization γ^* , in reference with q_1 . An illustration of $|V_{q_1(t) \rightarrow q_2^*(t)}|$ between points of two corresponding symmetrical curves is shown in Fig. 3.2. Here $|\cdot|$ denotes the \mathbb{R}^3 Euclidean Norm, and $|V_{q_1(t) \rightarrow q_2^*(t)}|$ represents the magnitude of $V_{q_1 \rightarrow q_2^*}$ on each t . A possible interpretation of this vector is the local deformations needed to go from a point $q_1(t)$ of the parameterized curve q_1 to the corresponding point $q_2^*(t)$

of the symmetrical curve q_2^* . We note that, in general, $|V_{q_1(t) \rightarrow q_2^*(t)}|$ and $|V_{q_2(t) \rightarrow q_1^*(t)}|$ are different ($|V_{q_1(t) \rightarrow q_2^*(t)}| \neq |V_{q_2(t) \rightarrow q_1^*(t)}|$).

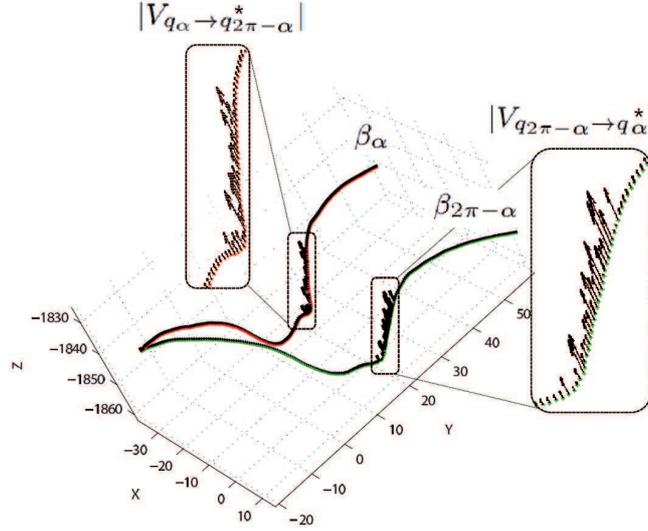


Figure 3.2 – 3D-Symmetry description from pairwise radial curves. Illustration of the deformation needed from β_α to fit $\beta_{2\pi-\alpha}$ ($|V_{q_\alpha \rightarrow q_{2\pi-\alpha}^*}|$) and inversely ($|V_{q_{2\pi-\alpha} \rightarrow q_\alpha^*}|$), on each point of the curve parameterized by t . q_α and $q_{2\pi-\alpha}$ denote the SRVs of bilateral symmetrical curves β_α and $\beta_{2\pi-\alpha}$, respectively, extracted from an arbitrary face.

3.4 GEOMETRIC FEATURE EXTRACTION

3.4.1 Feature extraction based on the morphological observations

Based on the mathematical framework described above, we extract four different types of 3D face descriptions. Each description reflects a high-level morphology cue in the face introduced at the beginning of this chapter. The descriptions are illustrated in Fig. 3.3. In each panel of this figure, the left part illustrates the extracted radial curves and the curve comparison strategy, the right part shows the extracted features as color-map on each point of the face. On each face point, the warmer the color, the lower the deformation magnitude. The definitions of these descriptions are as follows.

- **The 3D-Symmetry Description (3D-sym.)** shown in Fig. 3.3 (a) captures densely the deformation between bilateral symmetrical curves (β_α^S and $\beta_{2\pi-\alpha}^S$). It demonstrates that local asymmetries emerge

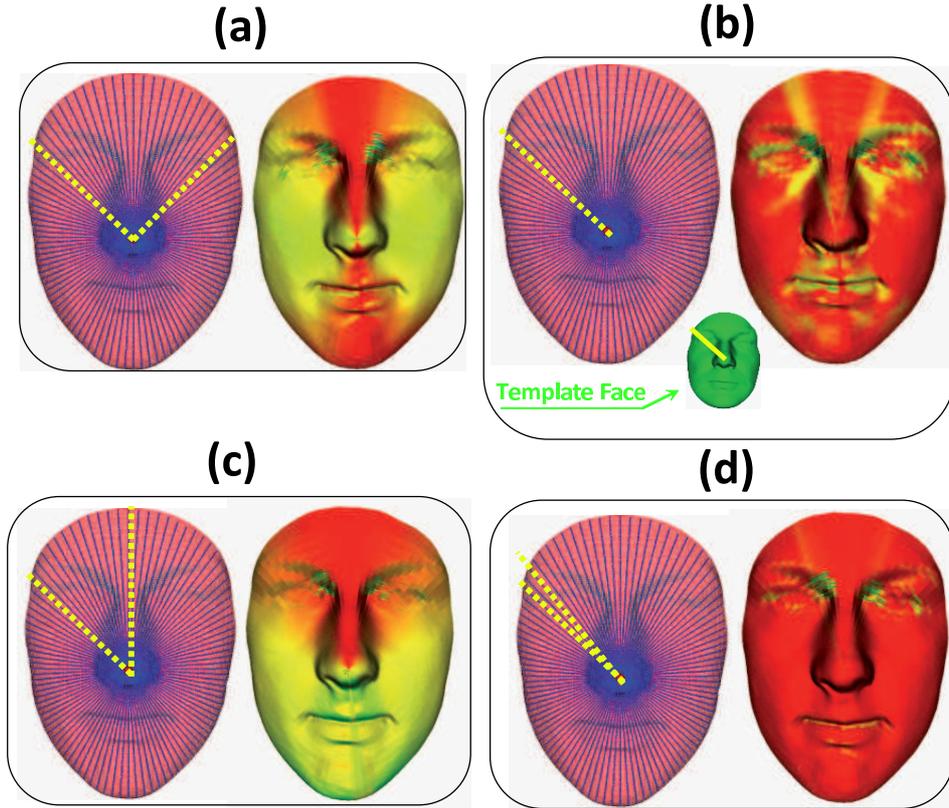


Figure 3.3 – Illustrations of different DSF features on 3D shape of the face S . (a) 3D-Symmetry Description: shape differences from each radial curve β_α^S to its symmetrical curve $\beta_{2\pi-\alpha}^S$; (b) 3D-Averageness Description: shape differences from radial curve β_α^S in a preprocessed face to radial curve β_α^T in face template (with same angle index α); (c) 3D-Spatial Description: shape differences from radial curve β_α^S to the middle-up radial curve β_0^S in the forehead; (d) 3D-Gradient Description: shape differences from radial curve β_α^S to its neighbor curve $\beta_{\alpha+\Delta\alpha}^S$

around the eyes, mouth, and the nose, and further away the middle-up facial curve, increasing asymmetry is observable. The 3D-sym. descriptor densely characterizes the bilateral asymmetry of the face. Considering that the face symmetry is closely related to Gender, Ethnicity and Age [72, 89, 16], this description allows to study the relationship between the bilateral symmetry and the 3D facial soft-biometrics.

- **The 3D-Averageness Description (3D-avg.)** shown in Fig. 3.3 (b) compares a pair of curves with the same angle index α from face β_α^S and an average template face β_α^T . The average face template T (presented in Fig. 3.4 and also in Fig. 3.3(b)) is defined as the middle point of the geodesic path which connects a male face (ID:

02463d548; Age: 48; White) to a female face (ID: 04200d74; Age: 21; White) arbitrarily selected from FRGCv2 dataset. With the radial curves representation, the pair-wise geodesic path between corresponding curves of the two scans are computed. Then the middle point of the geodesic is picked out by interpolation as the face template. As shown in Fig. 3.3(b), the averageness description shows that face shape differs mainly around the forehead, the eyes, the nose, and the mouth regions (yellow-green colors), to the mean face template. The 3D-Averageness description is a way to test the idea that different population groups deviate differently from a given mean face shape, as shown in previous anthropometry studies [10, 106, 1, 92, 69, 4, 6, 26, 53, 20, 5, 60, 105, 25]. Faces of different morphology should show different deformations to reach the template face.

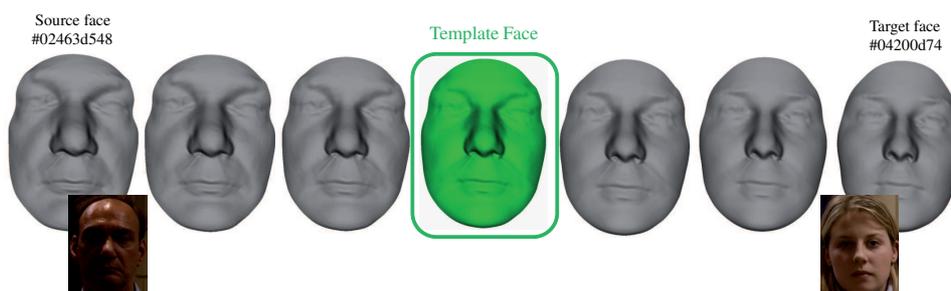


Figure 3.4 – The average face template T (presented also in Fig. 3.3(b)) is defined as the middle point of the geodesic path which connects a male face (ID: 02463d548; Age: 48; White) to a female face (ID: 04200d74; Age: 21; White).

- **The 3D-Spatial Description (3D-spat.)** shown in Fig. 3.3 (c) captures the deformation of a curve β_α to the middle-up curve β_0 , emanating from the nose tip and goes over the nose and the forehead in S . As shown in Fig. 3.3, this description shows a smooth deviation from the upper facial profile. As β_0 is the most rigid curve in the face, the 3D-Spatial description captures the intrinsic changes from the most rigid part of the face. It signifies the developmental differences of spacial facial features in comparison to the middle-up curve, which relate closely with Gender, Ethnicity and Age [60, 105, 90, 65].
- **The 3D-Gradient Description (3D-grad.)** shown in Fig. 3.3 (d) cap-

tures the differences between pairwise neighboring curves (β_α^S and $\beta_{\alpha+\Delta\alpha}^S$). The 3D-Gradient description captures the local deformations on the face, such as the facial wrinkles. It shows the smoothness and consistence in local face surface. As shown in Fig. 3.3 (d), it highlights some details around the eyes and the mouth (yellow-green colors). The *3D-grad.* can be viewed as a gradient operator over the face and can detect the wrinkles. As analyzed previously, this description is quite informative for face Gender, Ethnicity and Age perception .

From the definitions and analysis above, our DSF features derived from four types of high-level facial morphology cues are closely related to Gender, Ethnicity and Age in the face. In other words, in theory, they should be informative for face Gender, Ethnicity and Age recognition. Thus in the next, we examine the relationship between the facial features and the facial soft-biometric, to reveal what kind of information has been exactly obtained using these features.

3.4.2 Correlation between features and the facial soft-biometrics

The four descriptions above allow us to capture densely different perspectives of the facial shapes. With these perspectives, what clues can we perceive exactly from the face for each facial soft-biometric? In Figure 3.5, we show the magnitude of the correlation between the facial features and the facial soft-biometrics as colormap on the face. The correlations are calculated with the 466 earliest scans of FRGCv2 dataset. Formally, for all the 466 scans, the *ith* feature of a face description makes a one dimensional vector $D_i = (d_i^1, d_i^2, \dots, d_i^{466})$, where d_i^m denotes the *ith* feature of the *mth* face. Then for a facial biometric, the labels for 466 scans make a vector $L_a = (L_1, L_2, \dots, L_{466})$, where L_m denotes the attribute label of the *mth* face. Then the correlation between the *ith* feature of the description and the concerned attribute is given by the *Pearson Correlation Coefficient* between D_i and L_a . This coefficient is defined as the covariance of the two variables divided by the product of their standard deviations. For two variables X and Y , the *Pearson correlation coefficient* $p_{X,Y}$ is defined as $p_{X,Y} = \mathbf{cov}(X, Y) / (\sigma_X \sigma_Y)$,

where cov denotes the covariance, σ denotes the standard deviation. The absolute value of *Pearson Correlation Coefficient* ranges from 0 to 1. The higher the absolute value, the higher the linear correlation between the two variables. In our case, the higher the correlation between a feature and an attribute, the more informative is this feature for discriminating the concerned attribute.

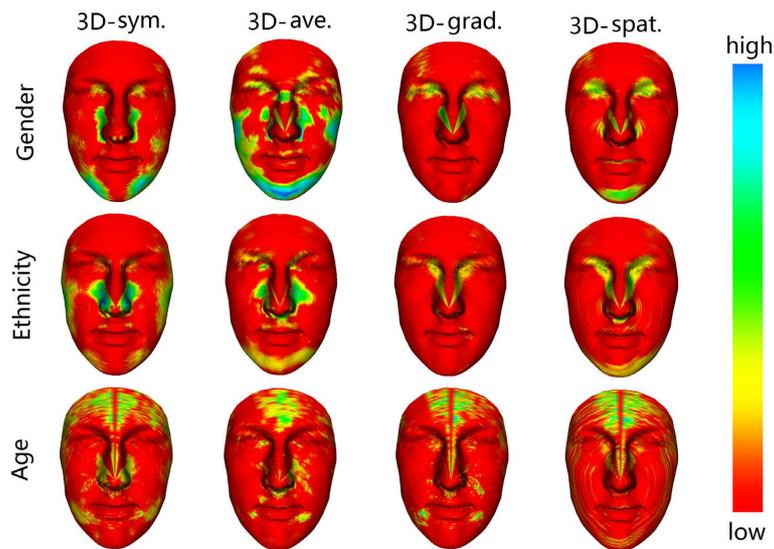


Figure 3.5 – Correlation between facial geometric features and facial soft-biometrics on each point of the face surface.

In Figure 3.5, the correlation between each description and each facial attribute (gender, age and ethnicity) is shown as colormap on each point of the face. Each column of the subfigures show the correlation between a face description and each facial attribute, and each row of subfigures show the correlation between each face description and a facial attribute. Warm colors indicate low correlation between a facial feature and an attribute, and cold colors indicate high correlation. With Figure 3.5, we can confirm three things.

First, the green and blue colors in these subfigures show that the face features are considerably correlated with the facial soft-biometrics. Thus, we confirm that the 3D face shape is informative for Gender, Ethnicity and Age.

Secondly, we confirm that the four descriptions give different and complimentary perspectives for perception of Gender, Ethnicity and Age.

They act as four different types of 'eyes' in face perception. The four descriptions 'see' different and complimentary clues of Gender, Ethnicity and Age in the 3D face.

For gender, the *3D-sym.* description 'sees' the inner eye corners, the connection area of the nose and the cheeks, and the chin-side regions. The *3D-ave.* description 'looks' at the eye-brows, the eyes, the nose, the lips, and gives big attention on the cheek-sides and the chin. The *3D-grad.* description emphasizes the eyes and the dorsal nasals of the nose, while the *3D-spat.* description also considers the chin and the sides of the nostrils. For Ethnicity, the *3D-sym.* description 'sees' the nose regions and the cheeks. The *3D-ave.* description 'looks' at the connection area of the nose and the cheek-sides, and the chin region. The *3D-grad.* description emphasizes the inner eye-corners and the dorsal nasals of the nose, while the *3D-spat.* description also emphasizes the chin and the area around the nostrils. For Age, the *3D-sym.* description 'sees' the whole forehead, the nose, the outer eye corners, and the the regions besides the mouth corners. The *3D-ave.* description mainly 'looks' at the center part of the forehead, the inner eye corners, and the nose surrounding regions. The *3D-grad.* description emphasizes the center part of the forehead, the eye corners, the nose bridge, and the mouth corners, while the *3D-spat.* description emphasizes the center part of the forehead, the eye corners, and the nose bridge.

Thirdly, it reveals that Gender, Ethnicity and Age information distributes in different regions of the 3D face. For Gender, the eyes, nose, cheek-sides, lips and the chin are particularly informative. It matches with the previous findings of sexual dimorphism in [1, 92], which claim that males have protuberant nose, eyebrows, chin and jaws than females, and the distance between top-lip and nose-base is longer. For Ethnicity, the eyes, nose, cheek-sides and chin are more informative. This echos the findings in [69, 4] which find that the Non-Asians have broader faces and noses, farther apart eyes, and lower fetal frontomaxillary facial angle (FMFA) measurements than Asians. For both Gender and Ethnicity, the forehead gives little information. While for Age, the forehead, together

with the nose, the eye corners and the mouth corners, show the strongest hints. It goes with the public knowledge that wrinkles usually develop in forehead, eyes, nose and mouth regions with age.

3.4.3 Statistical characteristics of the extracted features

The previous section have illustrated the correlation between the facial features and the facial attributes on each spatial point of the face. In this section, we propose to look at the statistical distributions of the facial DSF features among faces, to reveal the correlation between the facial features and the facial soft-biometrics in another way. In Fig. 3.6, Fig. 3.7, Fig. 3.8, we show the statistical distributions of the DSF features in the faces for different Gender (Male and Female), Ethnicity (Asian and Non-Asian) and Age groups (> 22 years old and < 23 years old). All the figures are generated with the DSF features of the 466 earliest scans of the 3D FRGCv2 dataset. In each panel of the figures, the x-axis shows the value of the concerning DSF features, and the y-axis shows the probability density of that value.

In Fig. 3.6, we show the value distribution of the DSF features for different Gender group. For the Averageness DSF, as shown in the first row of Fig. 3.6, clear difference is demonstrated between the feature distributions of male and female faces. The female faces have more high-value features (> 0.2) than male faces. It means that the female faces generally need more deformation to reach the face template. For the Symmetry DSF, as shown in the second row of Fig. 3.6, more high-value features (> 1.4) are found in the female faces, than in the male faces. It means that female faces have more high magnitude asymmetries than male faces. This observation matches the anatomical studies in [29, 27, 31] that the significant asymmetries between the two halves of the faces are greater in female faces than in male faces of corresponding age. In the third row of Fig. 3.6, we observe that with the Spatial DSF description, the male faces have more high-value features (> 0.15) than female faces. This implies that male faces present more global changes than female faces. This observation meets the anthropometry finding that the male faces features

are usually more prominent than female faces [106, 1, 92, 53]. In the last row of Fig. 3.6, the female faces show more high-value features (> 1.3) with the Gradient description. This observation goes with the common knowledge that the males have thicker and firmer skins than a woman's (because of higher level of testosterone hormones in their blood), and thus have higher resistance to skin deterioration (lines, wrinkles, skin laxity, etc.)¹. In summary of Fig. 3.6, as the male and female faces have shown significantly different distributions of feature values, it demonstrates that the proposed DSF features capture strong information for the discrimination between male and female faces.

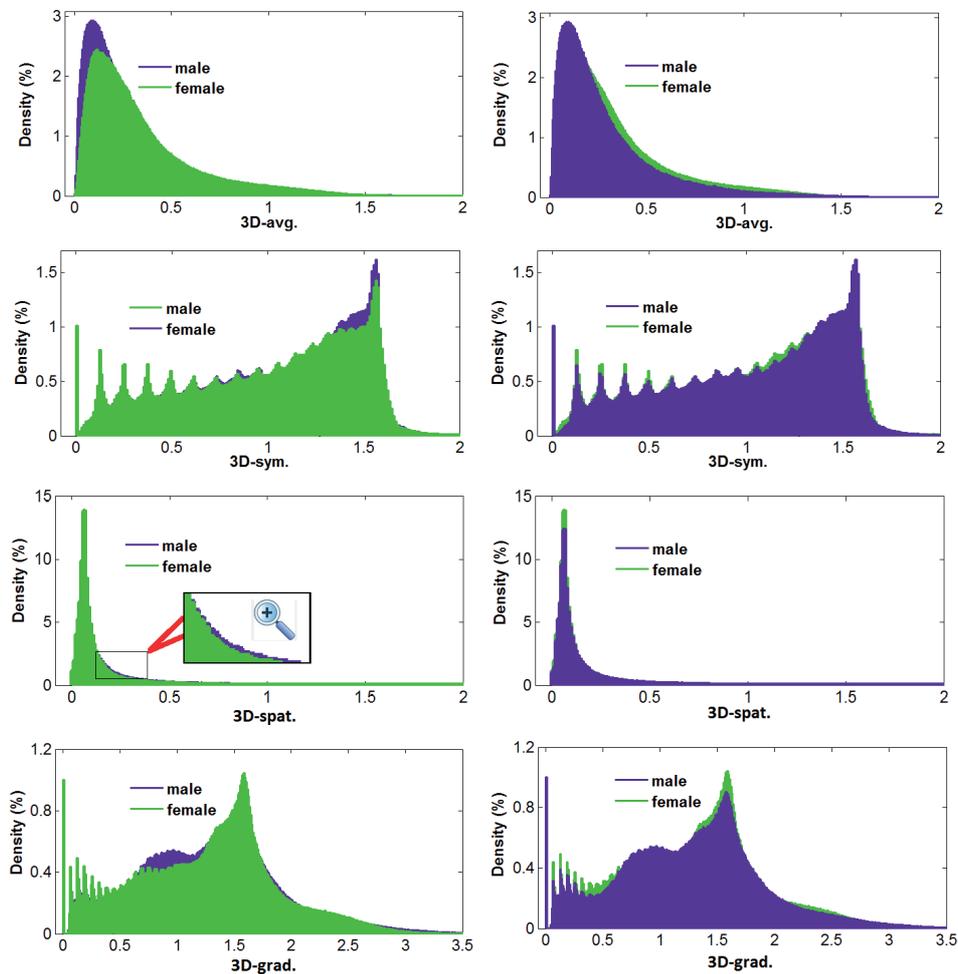


Figure 3.6 – Distribution of the DSF features considering different Gender groups.

The statistical distributions of the DSF features for different Ethnicity group are shown in Fig. 3.7. As shown in the first row of Fig. 3.7, with the

¹<http://www.menscience.com/blog/2008/04/men-have-thicker-skin-than-women.html>

Averageness DSF, the Non-Asian faces possess more high-value features (> 0.25) than Asian faces. For the Symmetry DSF, shown in the second row of Fig. 3.7, it is clear that Non-Asian faces have more high-value features (> 1.45) than the Asian faces. It means that statistically the Asian faces are less asymmetric than the Non-Asian faces. For the Spatial DSF, as shown in the third row of Fig. 3.7, the Non-Asian faces have more high-value features (> 0.05) than the Asian faces. It means that the Asian faces have less global shape changes than the Non-Asian faces. In the last row of Fig. 3.7, the distribution of the Gradient DSF shows that the Asian faces have more high-value features (> 0.6) than the Non-Asian faces. It means that the Asian faces have more local changes than the Non-Asian faces. In summary of Fig. 3.7, as the Asian and Non-Asian faces have shown significantly different distributions of feature values, it demonstrates that the proposed DSF features capture strong information for the discrimination between Asian and Non-Asian faces.

The distributions of the DSF features for different Age group (old groups > 22 and young group < 23) are shown in Fig. 3.8. As shown in the first row of Fig. 3.8, for the Averageness DSF, the old group has more high-value features (> 0.25) than the young group. It means the older the age, the more deformation is needed to reach the template face. For face Symmetry DSF, as shown in the second row of Fig. 3.8, the old group has more high-value features (> 1.45) than the young group. It means that face aging is associated with the increase of facial asymmetry. For the Spatial DSF, shown in the third row of Fig. 3.8, more high-value features (> 0.05) are found in the old group. It means that when face ages, the spatial difference also increases. The last row of Fig. 3.8 shows the distribution of the Gradient DSF. The old faces have less high-value features than the young faces. In summary of Fig. 3.8, as the faces from the young and old groups have shown significantly different distributions of feature values, it demonstrates that the proposed DSF features capture strong information revealing the age.

In summary, we find that the distributions of the DSF features have significant difference when considering different demographic population

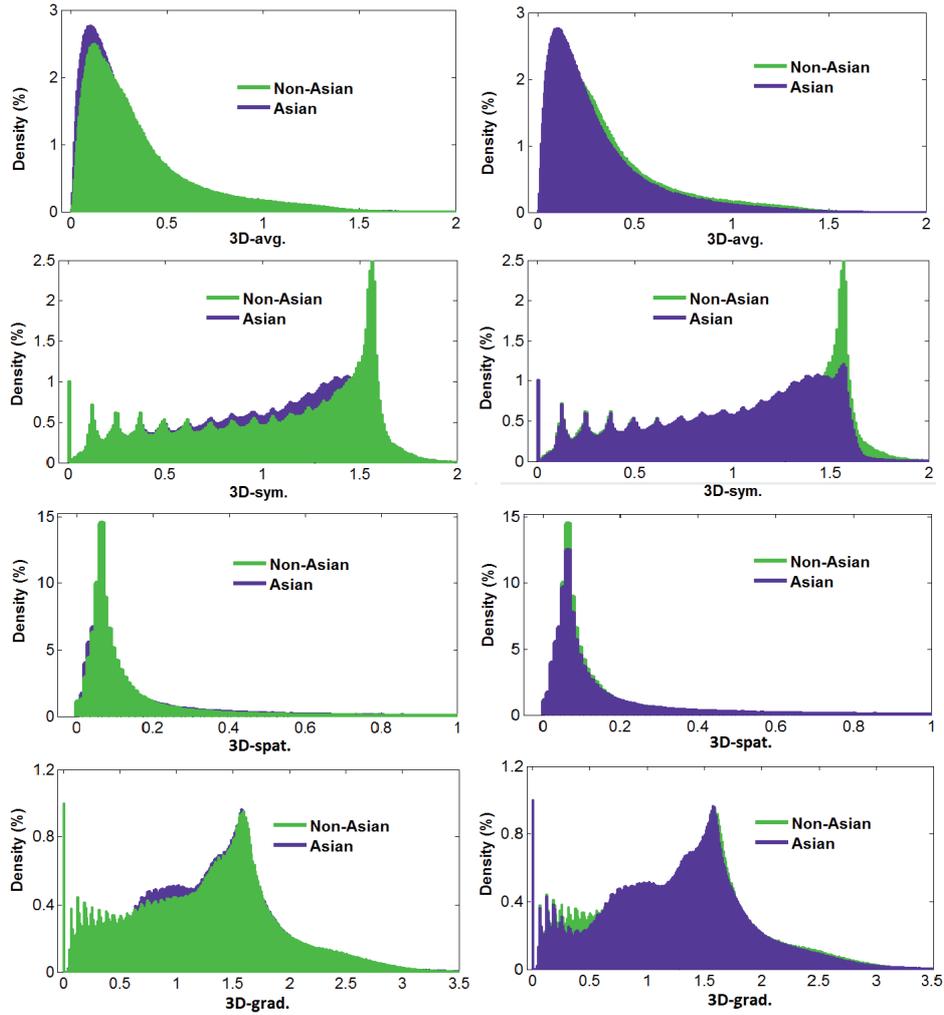


Figure 3.7 – Distribution of the DSF features considering different Ethnicity groups.

groups. It demonstrates that the proposed DSF features convey rich information for discriminating the concerning facial soft-biometrics. In addition, as different DSF features show significantly different pattern of distributions, it confirms that the four types of DSF features captures different morphology cues of the face.

3.5 CONCLUSIONS

In this chapter, we presented our methodology for extracting geometric features from 3D faces. The extracted Dense Scalar Field (DSF) features grounding on Riemannian shape analysis of facial radial curves are capable to capture densely the pair-wise shape differences between parametric

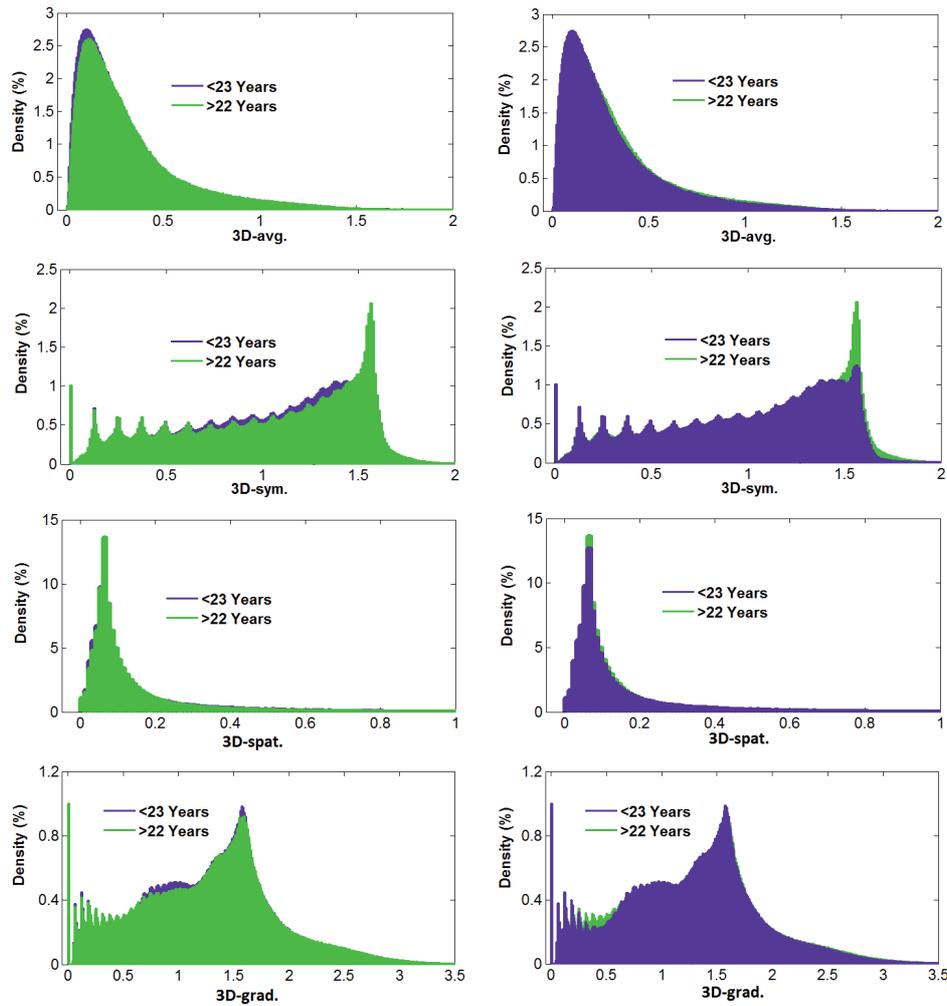


Figure 3.8 – Distribution of the DSF features considering different Age groups.

terized facial curves. With four high-level facial morphology cues which relate closely to Gender, Ethnicity and Age, we developed four types of geometric features, namely the face Averageness DSF, the Symmetry DSF, the Spatial DSF and the Gradient DSF. By examining the correlation between the facial features and facial soft-biometrics on each point of the facial surface, as shown in Fig. 3.5, we found that the proposed features have strong relevance to Gender, Ethnicity and Age. By analyzing the statistical distribution of these DSF features concerning different demographic groups, as shown in Fig. 3.6, Fig. 3.7 and Fig. 3.8, we confirmed that the proposed DSF features convey rich information for discriminating these soft-biometrics. In addition, these analysis revealed the relationship between the underlying morphological cues and the facial soft-biometrics.

For example, it reveals that the female/Non-Asian/old faces have more significant asymmetries than the male/Asian/young faces.

In the following chapter, we explore the usage of the proposed geometric features in automatic facial soft-biometric recognition. The experimental settings and the analysis will also be presented.

FACIAL SOFT-BIOMETRICS

RECOGNITION

SOMMAIRE

4.1	INTRODUCTION	46
4.2	EXPERIMENT SETTINGS	46
4.2.1	Evaluation Dataset and Protocols	46
4.2.2	Random Forest classifier/regressor	47
4.2.3	Dimensionality reduction methods	48
4.2.4	Feature-Level Fusion method	50
4.3	EXPRESSION-DEPENDENT SOFT-BIOMETRIC RECOGNITION	51
4.3.1	Experiments with high dimensional DSF features	51
4.3.2	Experiments with feature dimensionality reduction	55
4.4	EXPRESSION-INDEPENDENT SOFT-BIOMETRIC RECOGNITION	61
4.4.1	PCA-based Expression-Independent experiments	62
4.4.2	CFS-based Expression-Independent Experiments	64
4.4.3	Summary	66
4.5	FEATURE LEVEL FUSION EXPERIMENTS	68
4.5.1	Expression-Dependent Experiments	69
4.5.2	Expression-Independent Experiments	70
4.5.3	Summary	72
4.6	ROBUSTNESS TO FACIAL VARIANTS	72
4.7	ROBUSTNESS TO TRAINING SIZE	75
4.8	COMPARISON WITH THE STATE-OF-THE-ART	77
4.9	CONCLUSION	79

4.1 INTRODUCTION

In this chapter, we will make a comprehensive study of the proposed DSF features for the problem of Gender, Ethnicity and Age recognition. Gender classification is to automatically label a query instance, a 3D face scan in our case, into Male or Female. For Ethnicity, we consider ethnicity groups, the Asian and the Non-Asian groups. The aim of ethnicity classification is to automatically label a query instance into its ethnicity group. For Age, we take it as a continuous problem. Given a query instance, our goal is to automatically label it with an exact age. Detailed analysis of the experimental results are also presented in this chapter.

4.2 EXPERIMENT SETTINGS

4.2.1 Evaluation Dataset and Protocols

Our experiments are carried out on the Face Recognition Grand Challenge 2.0 (FRGCv2) dataset [80]. The FRGCv2 dataset was collected by researchers from the University of Notre Dame and contains 4007 3D face scans of 466 subjects with differences in gender, Ethnicity, age and expression. For gender, there are 1848 scans of 203 female subjects and 2159 scans of 265 male subjects. The ages of subjects range from 18 to 70, with 92.5% in the 18-30 age group. When considering Ethnicity, there are 2554 scans of 319 White subjects, 1121 scans of 99 Asian subjects, 78 scans of 12 Asian-southern subjects, 16 scans of 1 Asian and Middle-east subject, 28 scans of 6 Black-or-African American subjects, 113 scans of 13 Hispanic subjects, and 97 scans of 16 subjects whose Ethnicity are unknown. About 60% of the faces have a neutral expression, and the others show expressions of disgust, happiness, sadness and surprise. All the scans in FRGCv2 are near-frontal.

For extraction of the DSF features, we extract 200 radial curves on each face and 100 indexed points on each curve. So, each face is densely represented by 200×100 points for feature extraction. Furtherly, for each DSF description, we have also 20,000 feature points, which corresponds to each point of the face. The average time consumed for extracting all

200 curves for each face is 1.048 seconds. And the total computation time (including preprocessing) for each scan is less than 1.5 seconds. All our programs are developed in C++ and performed on Intel Core i5 CPU 2.53 GHZ with 4Go of RAM. Thus, our feature extraction procedure is very efficient. With these descriptions, we conduct two types of experiments:

- **Expression-Dependent Experiment** uses the 466 earliest scans from FRGCv2 for training and testing. The majority of the scans in this subset have neutral expression. This data subset leads to a possible study of the facial attribute recognition when imposing a neutral expression. [9, 36] have explored this data subset for 3D gender classification.
- **Expression-Independent Experiment** is based on the whole 4007 scans of FRGCv2 (about 40% are expressive). This makes possible the study of facial attributes recognition when varying the facial expressions. The whole FRGCv2 dataset has been extensively used to test the robustness of face recognition algorithms against facial expressions [24]. In [104], the Ethnicity classification results on FRGCv2 dataset are influenced strongly by facial expressions.

We use the *Leave-One-Person-Out (LOPO)* cross-validation approach in all our experiments, where each time the scans of one subject are used for testing, and the scans of the remaining subjects are used for training. Thus, there are altogether 466 folds in the cross-validation. The experiments are conducted in a Subject-independent fashion. Each subject is tested equally only once. There are two reasons for choosing the LOPO strategy, the first is its similarity to real-world like applications, and the second is it allows training with a maximum number of subjects. To make comparison to related works, we conducted also experiments following the 10-fold cross validation protocol.

4.2.2 Random Forest classifier/regressor

We adopt the Random Forest [13] to make the evaluation of our face descriptions in facial soft-biometric recognition. Random Forest is an en-

semble learning method that grows many decision trees $t \in \{t_1, \dots, t_T\}$ considering an attribute. To estimate the attribute from a new instance represented as a feature vector, each tree gives a decision result and the forest does the overall estimation. In growing of each tree, two types of randomness are introduced. First, to make the training set, a number of N instances are sampled randomly with replacement from the original data. Then at each node of the tree, a constant number of m ($m \ll M$) variables are randomly selected, and the best split on these m variables is used to split the node. The process goes on until the resulted subsets of the node are totally purified in label. The performance of the forest depends on the correlation between any two trees, and the strength of each individual tree. The forest error rate increases when the correlation decreases, or the strength increases. Reducing m reduces both the correlation and the strength. Increasing it increases both. Thus, an optimal m is needed for the trade-off between the correlation and the strength. In Random Forest, the optimal value of m is found by using the oob-error rate (out-of-bag-error rate).

For making the overall decision, in classification work, the forest predicts the attribute with majority voting. The classification mode of Random Forest is designed for instances with discrete class labels, such as the Gender and Ethnicity labels. While in regression tasks, it takes the average of predictions. The regression mode of Random Forest is designed for instances with continuous class labels, such as the Age labels. Thus, in our work, we use Random Forest in classification mode for Gender and Ethnicity recognition, and in regression mode for Age estimation.

4.2.3 Dimensionality reduction methods

As the original DSF features have as many as 20,000 dimensions, we apply two types of dimensionality reduction methods to have a salient and compact representation of the original features. We propose to use two well-known feature dimensionality reduction techniques, the supervised Correlation-based Feature Selection (CFS) which picks out a salient subset of features from the original feature space, and the un-supervised Princi-

pal Component Analysis (PCA) which captures the main variance of the original features in a lower dimensional subspace by orthogonal transformation. For CFS, it needs the label information of the concerning attribute (soft-biometric) for each instance. While for PCA, it works naturally without using any label information of the instances.

Correlation-based Feature Selection (CFS)

Feature subset selection is the process of identifying and removing as much irrelevant and redundant information as possible [45]. There are mainly two types of feature selection methods, *the filter* methods which use heuristics based on general characteristics of the data to evaluate the merit of feature subsets, and *the wrapper* methods which use an induction algorithm along with a statistical re-sampling technique such as cross-validation to estimate the final accuracy of feature subsets [85]. We choose a filter method for feature selection, named the Correlation-based-Feature-Selection (CFS) [45], because the filters operate independently of learning algorithm and are generally much faster than wrappers. The chosen CFS filter comprises of two parts, a feature correlation measure using the Pearson's correlation coefficient, and a Best-First heuristic search algorithm which moves through the search space by greedy hill-climbing augmented with a back-tracking facility. In practice, we perform feature selection for all the three facial biometrics, the gender (labeled as male and female), the ethnicity (labeled as Asian and Non-Asian) and the age (labeled in two groups, more than 22 and less than 23). After Feature selection, we retain 200-400 features for each description. Thus, the feature selection procedure significantly reduces the size of the features.

Principal Component Analysis (PCA)

The Principal Component Analysis (PCA) [94] is a mathematical procedure using orthogonal transformation to convert a set of observations of variables into a set of linearly independent variables (principal components). This transformation is defined in such a way that the first principal component has the largest possible variance (that is, accounts for as

much of the variability in the data as possible), and each succeeding component in turn has the highest variance possible under the constraint that it is orthogonal to the preceding components. By selecting a number of principal components, feature dimensionality reduction is achieved. Compared to supervised feature dimensionality reduction, such as the CFS, one advantage of PCA is the independence to the labels of the instances (supervision information). One disadvantage of PCA is the loss of the spatial correspondence of features and the studied subject as the features are transformed into another space.

4.2.4 Feature-Level Fusion method

When multi types of features are involved, fusion method is a common solution for machine learning problems. In the literature, two types of fusion method exist. The first is named *early fusion* or *feature-level fusion*, which combines the multi-modality features into a single representation, such as the concatenation of the features. *Early fusion* yields really a multi-modality representation, as the features are integrated from the beginning. Another advantage is it only requires only one training phase. The counterpart is named *late fusion* or *decision-level fusion*, which requires separate supervised learning stage for each type of feature, and then fuses the machine learning scores to evaluate semantic concepts. For example, the probabilistic models learned from each feature could be combined to yield a final decision. In our case, we have four types of DSF features derived from high-level semantic concepts. Our goal is to demonstrate whether the four descriptions are complimentary or not in facial soft-biometric recognition tasks. The *early fusion* will serve better our goal, as it contains completely the information from each description. Also, as stated, compared with *late fusion*, the *early fusion* needs only one training phase and do not need separate supervised learning stage. Thus, in our work, we choose the *early fusion*, which combines our four types of DSF features through concatenation.

4.3 EXPRESSION-DEPENDENT SOFT-BIOMETRIC RECOGNITION

In this section, we explore the original DSF features of the 466 earliest scans for facial attribute recognition. As described previously, Random Forest is used to evaluate the strength of our face descriptors in the recognition tasks. Considering the problem of high dimensionality in the original DSF features, we also test the performance of these features in combination of dimensionality reduction methods for soft-biometrics recognition.

4.3.1 Experiments with high dimensional DSF features

In this subsection, we perform LOPO experiments on the original DSF features extracted from the 466 earliest scans of FRGCv2 (Expression-Dependent), for Gender, Ethnicity and Age recognition. The experimental results are shown in Fig. 4.1. In each panel, the y-axis shows the recognition performance in the LOPO experiment. The x-axis shows the face descriptions.

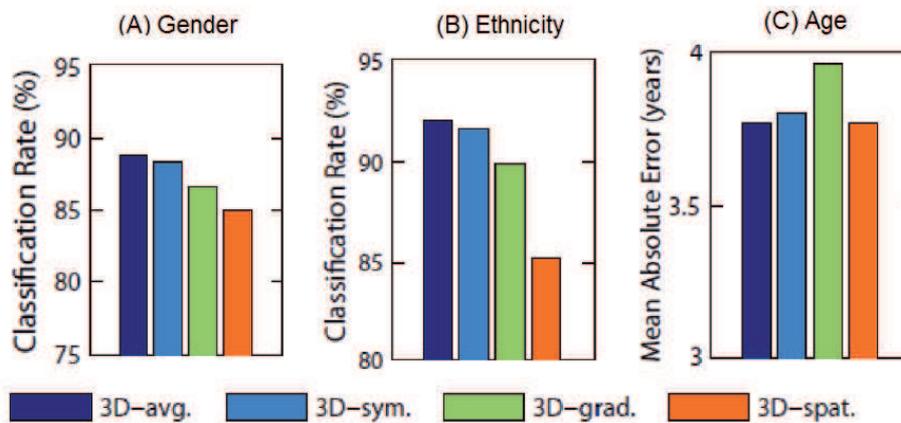


Figure 4.1 – Expression-Dependent Experiments with the original DSF features

Gender Classification

As shown in the left panel of Fig. 4.1, the Gender classification rate is generally $> 80\%$ for each DSF description. It demonstrates that the face descriptors are able to capture the sexual dimorphism in face.

Considering the results in each description, with the $3D-avg.$ descriptor, we achieve 89.06% Gender classification rate. It confirms the studies on sexual dimorphism [106, 1, 92, 53] which claim that Male and Female faces present statistically different morphological features in the shape. With the $3D-sym.$ descriptor, we achieve 87.77% Gender classification rate. It confirms that facial asymmetry is related to the Gender [73]. With the $3D-grad.$ descriptor, we achieve 87.12% Gender classification rate. With the $3D-spat.$ descriptor, we achieve 84.98% Gender classification rate. Results from $3D-grad.$ and $3D-spat.$ confirm that sexual dimorphism exists both locally and globally in face shape. These results are detailed in confusion matrices in Tables 4.1-4.4. For each description, the recognition rate for females and males are very close. It means that our approach is not biased to specific Gender. Slightly higher results are usually found for males. This difference is probably due to the fact that more male faces (263) are available for training than female scans (203).

Table 4.1 – confusion matrix of $3D-avg.$

	female	male
female	86.70%	13.30%
male	9.13%	90.87%
<i>Recognition rate =88.84%</i>		

Table 4.2 – confusion matrix of $3D-sym.$

	female	male
female	86.70%	13.30%
male	11.41%	88.59%
<i>Recognition rate =87.77%</i>		

Table 4.3 – confusion matrix of $3D-grad.$

	female	male
female	82.76%	17.24%
male	12.17%	87.83%
<i>Recognition rate =85.62%</i>		

Table 4.4 – confusion matrix of $3D-spat.$

	female	male
female	76.85%	23.15%
male	10.27%	89.73%
<i>Recognition rate =84.12%</i>		

Ethnicity Classification

As shown in the middle panel of Fig. 4.1, the Ethnicity classification rate is $> 85\%$ for each DSF description. These results demonstrate that the face descriptors are able to capture the discriminative information of Ethnicity in 3D face. Considering the results in each description, with the $3D-avg.$ descriptor, we achieve 92.06% Ethnicity classification rate. It confirms the findings of previous studies [69, 4, 6] that a significant difference exists between asian and Non-Asian faces. With the $3D-sym.$ descriptor, we

achieve 91.42% Ethnicity classification rate. It shows that the bilateral asymmetry can play an important role in Ethnicity classification. With the *3D-grad.* descriptor, we achieve 87.12% Ethnicity classification rate. With the *3D-spat.* descriptor, we achieve 84.33% Ethnicity classification rate. Results from *3D-grad.* and *3D-spat.* show that discriminative cues of Ethnicity exist both locally and globally in face shape. The details of these results are shown in confusion matrices in Tables 4.5-4.8. For each description, the recognition rate for Asian faces are significantly lower than Non-Asian faces. This difference is probably due to significantly more Non-Asian faces (354) are available for training than the Asian faces (112).

Table 4.5 – confusion matrix of *3D-avg.*

	Asian	Non-Asian
Asian	71.43%	28.57%
Non-Asian	1.41%	98.59%
<i>Recognition rate =92.06%</i>		

Table 4.6 – confusion matrix of *3D-sym.*

	Asian	Non-Asian
Asian	78.57%	21.43%
Non-Asian	4.52%	95.48%
<i>Recognition rate =91.42%</i>		

Table 4.7 – confusion matrix of *3D-grad.*

	Asian	Non-Asian
Asian	50.89%	49.11%
Non-Asian	1.41%	98.59%
<i>Recognition rate =87.12%</i>		

Table 4.8 – confusion matrix of *3D-spat.*

	Asian	Non-Asian
Asian	76.35%	23.65%
Non-Asian	9.51%	90.49%
<i>Recognition rate =84.33%</i>		

Age Estimation

The age estimation accuracy is typically measured by the mean absolute error (MAE) and the cumulative score (CS). The MAE is defined as the average of the absolute errors between the estimated age and the ground truth age, while the CS, proposed firstly by [34] in age estimation, shows the percentage of cases among the testing set where the absolute age estimation error is less than a threshold. As shown in the right panel of Fig. 4.1, the MAEs for all the descriptions are always < 4 years. These results demonstrate that the face descriptors are able to capture the aging patterns in 3D face. Following the literature, we show in Table 4.9 the age estimation accuracy in each age group for each description. For all the descriptors, the age estimation accuracy is always higher in young age groups, than in old age groups. Considering the number of training

subjects, as shown in the last column of the table, we assume that the big decrease of the number of scans in aged groups (from about 200 to about 20) accounts largely for this performance decline. With the *3D-avg.* description, we achieve 3.76 years MAE. It demonstrates that, statistically, face shapes differ with age. With the *3D-sym.* description, we achieve 3.79 years MAE. It confirms the idea that faces of different age have different magnitude of asymmetry. With the *3D-grad.* description, we achieve 3.94 years MAE. With the *3D-spat.* descriptor, we achieve 3.76 years MAE. Results from *3D-grad.* and *3D-spat.* indicate that face aging is both a local and a global process in face shape.

Table 4.9 – Age estimation details with the DSF features

	3D-avg.	3D-sym.	3D-grad.	3D-spat.	Fusion	# scans
≤ 20	3.48	3.43	3.77	3.30	3.93	185
(20,30]	2.18	2.58	2.32	2.38	2.29	246
(30,40]	9.99	7.60	10.05	8.92	7.03	20
≥ 40	24.82	23.66	24.56	25.36	24.45	15
All	3.76	3.79	3.94	3.76	3.63	466

Another perspective for the age estimation results is the Cumulative Score (CS). Fig. 4.2 shows the Cumulative Scores for the four descriptors with Random Forest. In Fig. 4.2, the x-axis is the level of Mean Absolute Error, the y-axis shows the cumulative score of accuracy by percentage of acceptance. Thus, a point (a,b) on the curve shows, with a Mean Absolute Error tolerance of a years, it achieves an acceptance of b percent. From Figure 4.2, we observe that, with an Error Level of 5 years, we achieve an acceptance of more than 75% over the 466 scans; when the Error Level is 10 years, the cumulative score increases to more than 90%¹.

Summary

The experiments above have made a comprehensive evaluation of the strength of our DSF-based descriptions in facial soft-biometric recognition. Experimental results confirm that discriminating cues of Gender, Ethnicity and Age exist in 3D shape of face, in terms of averageness, symmetry, gradient and spatiality. Results also reveal that the four DSF-based

¹To the best of our knowledge, this is the first work in the literature which addresses the problem of age estimation using 3D facial shapes.

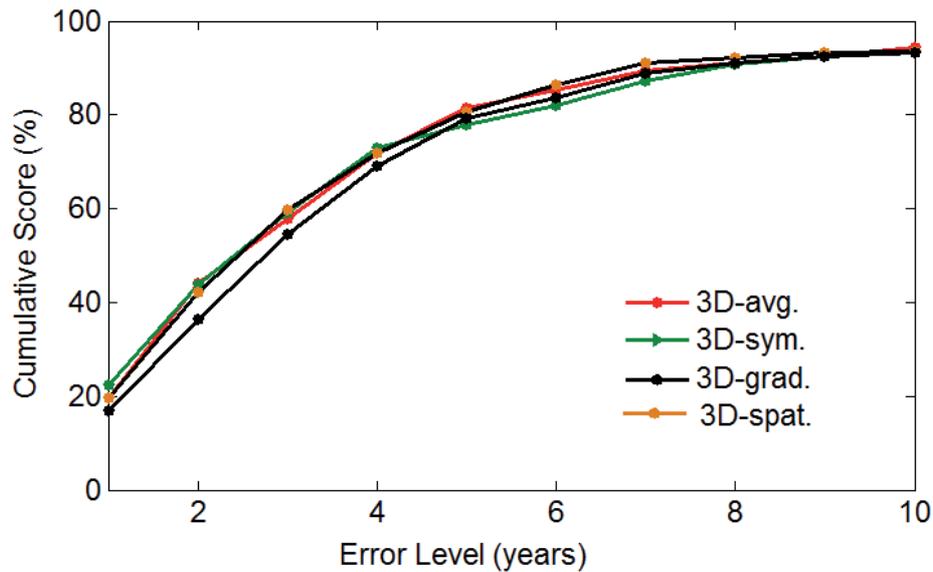


Figure 4.2 – Cumulative Score of age estimation for different DSF features

descriptions are capable to capture the discriminating cues of these soft-biometrics in 3D shape. With the fusion method, we always achieve better recognition performance than with individual description. This demonstrate that the proposed DSF descriptions are complimentary to each other in revealing the cues of Gender, Ethnicity and Age in 3D face. With the fusion, we achieve 90.12% gender classification rate, 93.56% ethnicity classification rate, and 3.63 years age estimation MAE, under the LOPO protocol.

4.3.2 Experiments with feature dimensionality reduction

Since the original DSF features have as many as 20,000 dimensions, we try two types of feature dimensionality reduction methods on the proposed DSF features, and then examine the performance of the resulted features in facial soft-biometric recognition. The two dimensionality reduction methods are the Principal Component Analysis (PCA) method and the Correlation-based Feature Selection method (CFS).

PCA-based Expression-Dependent Experiments

With the 466 earliest scans, we perform PCA on each type of DSF features, and explore the performance of facial soft-biometric recognition al-

gorithms with the number of principal components changing from 1 to 100. The recognition results in LOPO experiments with Random Forest classifier/regressor are shown in Fig. 4.3.

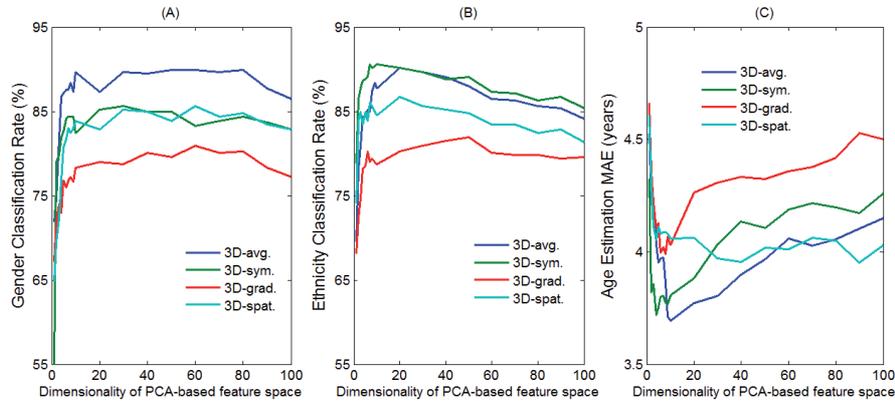


Figure 4.3 – *Expression-Dependent Experiments with PCA transformed features. In subplot (A) and subplot (B), the y-axis show the classification rate of Gender and Ethnicity, respectively. In subplot (C), the y-axis shows the Mean Absolute Error (MAE) of Age estimation.*

As shown in Fig. 4.3 (A), the Gender Classification rate for each description reaches $> 75\%$, when 5 principal components are used. When increasing the number of principal components, the Gender classification performances stay relatively stable. The performances decrease after 80 principal components, probably due to the fact that, thereafter, too much noises are introduced in the principal components. With the *3D-avg.* description, we achieve 89.70% Gender classification rate with 10 principal components. With the *3D-sym.* description, we achieve 85.62% using 20 principal components. With the *3D-grad.* description, we achieve 80.90% using 60 principal components. With the *3D-spat.* description, we achieve 85.62% using 60 principal components. Recall that with Random Forest classifier and the original DSFs, the gender classification rate reaches 88.84% in the *3D-avg.* description, 88.41% in the *3D-sym.* description, 85.62% in the *3D-grad.* description, and 84.12% in the *3D-spat.* description. Thus, the PCA-based features achieve comparable results with the original DSF features in the *3D-avg.* and *3D-spat.* descriptions, and lower results than the original DSF features in the *3D-sym.* and *3D-grad.* descriptions. While, considering that all the PCA-based features achieve 80% Gender classification rate, and the feature dimensionality changes from 20,000 to

10-60 after PCA transformation, we conclude that the PCA-based feature dimensionality reduction is useful in 3D Gender classification with our DSF-based descriptions.

For the Ethnicity Classification, as shown in Fig. 4.3 (B), the recognition rate for each description exceeds 80%, when 6 principal components are used. Similar to the Gender classification results, when increasing the number of principal components, the Ethnicity classification performances also stay relatively stable. Considering the results, With the *3D-avg.* description, we achieve 90.12% Ethnicity classification rate with 15 principal components. With the *3D-sym.* description, we achieve 90.55% using 20 principal components. With the *3D-grad.* description, we achieve 80.90% using 30 principal components. With the *3D-spat.* description, we achieve 86.69% using 15 principal components. Recall that with Random Forest classifier and the original DSFs, the Ethnicity classification rate reaches 92.02% in the *3D-avg.* description, 91.42% in the *3D-sym.* description, 87.12% in the *3D-grad.* description, and 84.33% in the *3D-spat.* description. Thus, the PCA-based features achieve lower results than the original DSF features except only in the *3D-spat.* description. While, considering that > 80% Ethnicity classification rate has been achieved with 15-20 dimension features (<<20000), the PCA-based feature dimensionality reduction is still viewed as helpful in 3D Ethnicity classification with our DSF-based descriptions.

For Age estimation, as shown in Fig. 4.3 (A), we achieve < 4% years MAE for each description, when 6 principal components are used. Except for the *3D-spat.*, all other descriptions achieve its lowest MAE with ≤ 10 PCA features. With the *3D-avg.* description, we achieve 3.71 years MAE with 10 principal components. With the *3D-sym.* description, we achieve 3.70 years MAE with 5 principal components. With the *3D-grad.* description, we achieve 4.13 years MAE with 5 principal components. With the *3D-spat.* description, we achieve 3.90 years MAE with 40 principal components. Recall that with Random Forest classifier and the original DSFs, the Age estimation MAE is 3.76 years in the *3D-avg.* description, 3.79 years in the *3D-sym.* description, 3.94% in the *3D-grad.* description, and 3.76

year in the *3D-spat.* description. Thus, the PCA-based features achieve comparable MAEs than the original DSF features. For the results in each age group, as shown in Table 4.10, the results are relatively lower in the old age group, than with the original DSF features. In terms of cumulative scores, as shown in Fig. 4.4, we achieve about 80% acceptance with a 5 years estimation error, and > 90% acceptance with a 10 years error. Thus, these results are comparable to the results with the original DSF features, in terms of the CS.

Table 4.10 – Age estimation details with PCA transformed features

	3D-avg.	3D-sym.	3D-grad.	3D-spat.	# scans
≤ 20	3.36	3.34	3.82	3.39	185
(20,30]	2.24	2.14	2.48	2.61	246
(30,40]	10.38	10.07	10.35	9.25	20
≥ 40	23.39	25.36	26.68	24.10	15
All	3.71	3.70	4.13	3.90	466

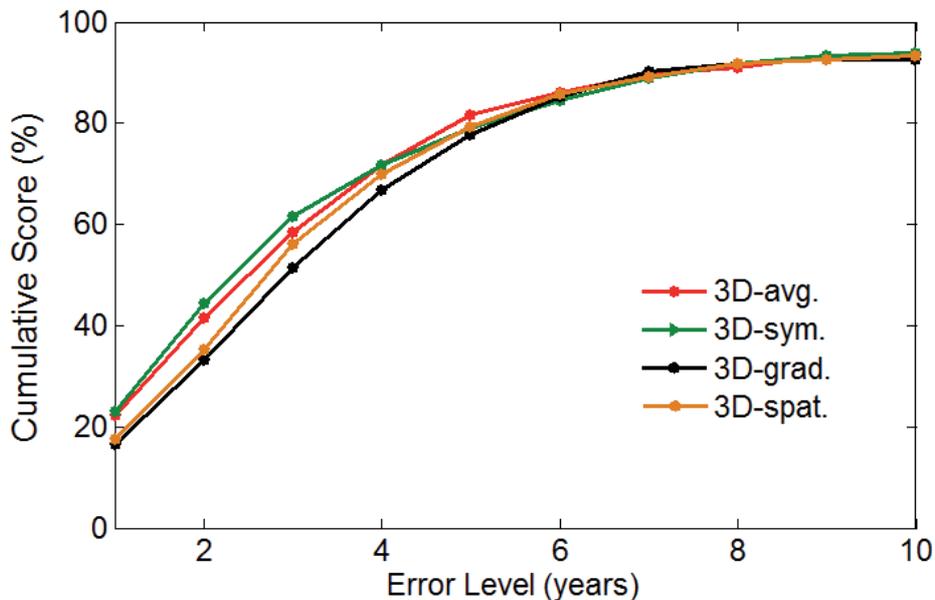


Figure 4.4 – Cumulative Score of age estimation with PCA transformed features

CFS-based Expression-Dependent Experiments

As described previously, the Correlation-based Feature Selection (CFS) is a supervised *filter* method for feature selection, which bases on the *Pearson Correlation Coefficient* between features and labels. After feature selection, a salient set of features which gives the highest overall correlation to the

interested labels will be kept. Compared with PCA, the CFS does not transform the original features, thus it makes the trace between the original features and the selected features possible. With the four types of DSF features, we perform feature selection on the 466 earliest scans of the FRGCv2 dataset for each facial soft-biometric. For Gender, we consider male and female groups for feature selection. For Ethnicity, we consider Asian and Non-Asian for feature selection. For age, we separate the 466 scans into old group (≥ 23 years) and young group (< 23 years) for feature selection. After feature selection, the dimensionality of the original DSF features is reduced from 20000 to a subset of 200-400 features. Thus, this procedure reduces significantly the size of the feature. After that, we perform the Expression-Dependent experiments with Random Forest to test the strength of these selected features. The corresponding Gender classification, Ethnicity classification and Age estimation results are shown in Fig. 4.5.

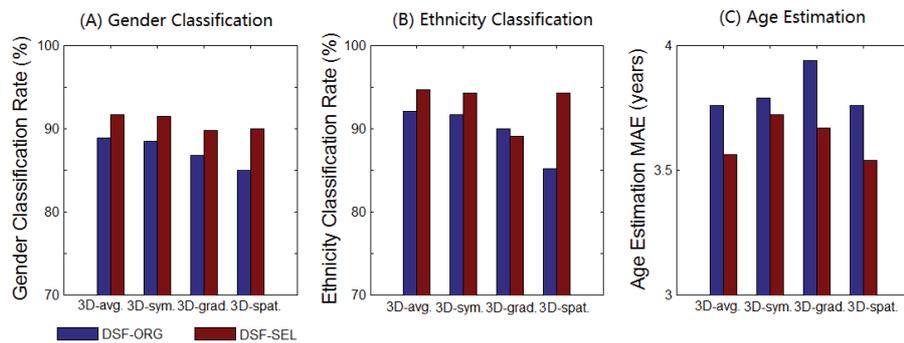


Figure 4.5 – Expression-Dependent Experiments with CSF selected features

In Fig. 4.5, for comparison purpose, the recognition results with the original DSF features on the 466 scans are also shown in each bar plot. These results are marked in blue color and labeled as *DSF-ORG* in the legends. For Gender classification, as shown in Fig. 4.5 (A), the performance with selected features (*DSF-SEL*) is always 3-5% higher than that with the original DSF features. It demonstrates that the feature selection procedure gives significant improvement to the gender classification performance. With the selected features of *3D-avg.* and *3D-sym.*, we achieve 91.63% and 91.42% Gender classification rate respectively. These results are also further higher than those with the transformed PCA features, as shown

in Fig. 4.3 (A). For Ethnicity classification, as shown in Fig. 4.5 (B), the performance with selected features (*DSF-SEL*) is always 2-7% higher than that with the original DSF features, except only for the *3D-grad.*, where the results are also comparable. Also, only except the *3D-grad.*, all other descriptions achieve more than 94% Ethnicity classification rate. Again, these results are further higher than the results from transformed PCA features in Fig. 4.3 (B). For Age estimation, the results shown in Fig. 4.5 (C) also demonstrate clear enhancement after feature selection. The MAEs are always lower compared with the original DSF features, and also lower than the PCA transformed features. When looking into the MAE in each age group, as shown in Table 4.11, the MAE in old groups, especially in the age range of (30,40], have been significantly improved than with the original DSF features. In terms of cumulative score, as shown in Fig. 4.6, the selected features achieve about 80% acceptance with an error level of 5 years, and > 90% acceptance with an error level of 10 years. These results are comparable with the results from the original DSF features. Thus, generally speaking, the feature selection method improves the age estimation results.

Table 4.11 – Age estimation details with CFS selected features

	3D-avg.	3D-sym.	3D-grad.	3D-spat.	# scans
≤ 20	4.68	4.37	4.61	3.46	185
(20,30]	2.13	2.40	2.24	2.35	246
(30,40]	7.32	5.99	6.95	6.42	20
≥ 40	22.90	25.10	23.75	22.92	15
All	3.56	3.72	3.67	3.54	466

Summary

In this subsection, we have explored the usage of the PCA and CFS for feature dimensionality reduction of the proposed DSF features in facial soft-biometric recognition tasks under the *Expression-Dependent* setting. Both the PCA and CFS have reduced significantly the dimensionality of the original DSF features, and achieved comparable results compared with the DSF features for all the three soft-biometrics. In comparison between the PCA and the CFS, the CFS has shown better performance than PCA, in terms of the Gender classification rate, the Ethnicity classification rate,

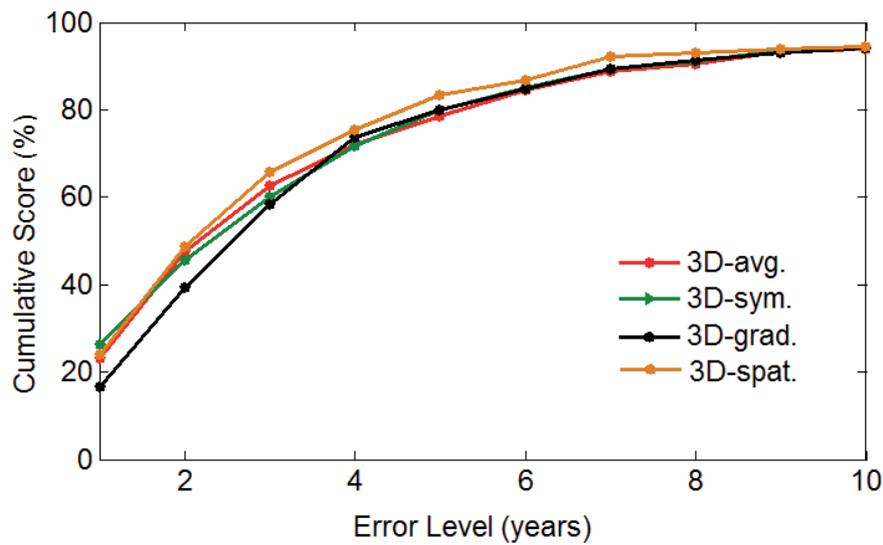


Figure 4.6 – Cumulative Score of age estimation with CFS selected features

and the Age estimation MAE/CS. The results from CFS always outperform the results with the original high dimensional DSF features for each face description and each facial soft-biometric.

4.4 EXPRESSION-INDEPENDENT SOFT-BIOMETRIC RECOGNITION

Facial expression change is one of the most difficult challenges in face analysis. The deformation of facial bones and facial soft tissues increases significantly the difficulty for face registration and comparison. As a result, the robustness to facial expression changes is one of the most important property of the related recognition algorithms. Thus, in this section, we test our facial soft-biometric approach on the whole FRGCv2 dataset, to reveal its robustness to facial expression changes. As previously stated, the high dimensionality of original DSF features has posed overwhelming challenge in computation cost for experimental evaluation on the whole FRGCv2 dataset. Beyond this, the recognition performance probably also suffers from the redundant and irrelevant feature dimensions. Thus, we explore again the usage of the unsupervised Principal Component Analysis (PCA) method and the supervised Correlation-based Feature Selection (CFS) method on the DSF features for feature dimensionality reduction, and then carry out the experiments for recognizing the facial soft-

biometrics. Here, we note again that the Random Forest is used in LOPO subject-independent cross validation for all the experiments.

4.4.1 PCA-based Expression-Independent experiments

In the Expression-Independent Experiments with the PCA transformed features, we perform PCA on each type of DSF features, and explore the performance of facial soft-biometric recognition algorithms with the number of principal components changing from 1 to 45. The recognition results in LOPO experiments with Random Forest classifier/regressor are shown in Fig. 4.7.

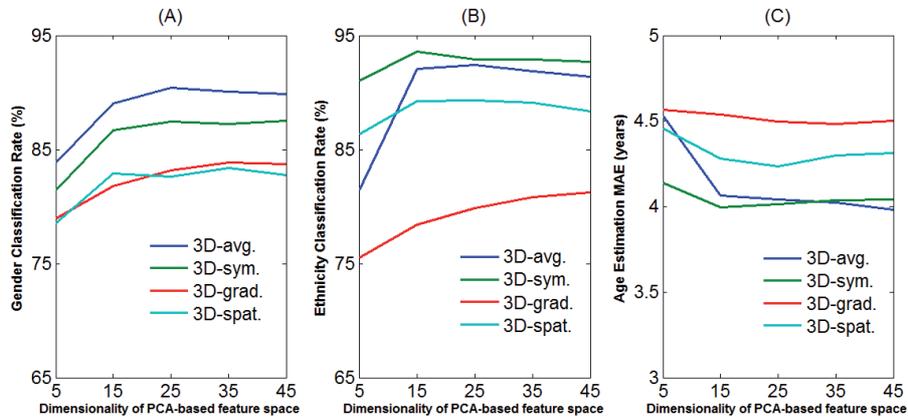


Figure 4.7 – Expression-Independent Experiments with PCA transformed features

As shown in Fig. 4.7 (A), the Gender Classification rate for each description exceeds 78%, when 5 principal components are used. When increasing the number of principal components, the Gender classification performances increase to $> 80\%$, and stay relatively stable after 15 components. Considering the achievements of results, with the *3D-avg.* description, we achieve 90.39% Gender classification rate with 25 principal components. With the *3D-sym.* description, we achieve 87.52% using 45 principal components. With the *3D-grad.* description, we achieve 83.90% using 35 principal components. With the *3D-spat.* description, we achieve 83.37% using 35 principal components. In Tables 4.12-4.15, we show the confusion matrix of these results. From these tables, we observe that the results for both gender are effective in each description (generally $> 80\%$), and the results for the male scans are slightly better than those for the

female scans. Considering that all the PCA-based features achieve 80% Gender classification rate, and the feature dimensionality changes from 20000 to 25-45 after PCA transformation, we conclude that the PCA-based feature dimensionality reduction is useful in 3D Gender classification with our DSF-based descriptions, even with the expression changes in the 3D faces.

Table 4.12 – confusion matrix of 3D-avg.

	female	male
female	89.33%	10.67%
male	8.72%	91.28%
<i>Recognition rate =90.38%</i>		

Table 4.13 – confusion matrix of 3D-sym.

	female	male
female	85.76%	14.24%
male	10.99%	89.01%
<i>Recognition rate =87.51%</i>		

Table 4.14 – confusion matrix of 3D-grad.

	female	male
female	81.48%	19.52%
male	14.05%	85.95%
<i>Recognition rate =83.89%</i>		

Table 4.15 – confusion matrix of 3D-spat.

	female	male
female	77.58%	22.42%
male	10.68%	88.32%
<i>Recognition rate = 83.37%</i>		

For the Ethnicity Classification, as shown in Fig. 4.7 (B), the recognition rate for each description exceeds 75%, when 5 principal components are used. When increasing the number of principal components, the Ethnicity classification performances stay effective and generally enhanced. With the 3D-avg. description, we achieve 92.40% Ethnicity classification rate with 25 principal components. With the 3D-sym. description, we achieve 93.55% using 15 principal components. With the 3D-grad. description, we achieve 78.40% using 15 principal components. With the 3D-spat. description, we achieve 88.36% using 45 principal components. In Tables 4.16-4.19, we show the confusion matrix of these results. From these tables, we observe that the results for the Non-Asian scans are significantly higher than those for the Asian scans, especially in the 3D-grad. and the 3D-spat. descriptions. It demonstrates that the imbalance of training data has great influence on the experimental results. Except for the 3D-grad. description, the other descriptions achieve > 85% Ethnicity classification rate, and the feature dimensionality changes from 20000 to 15-45 after PCA transformation. Thus, we conclude that the PCA-based feature dimensionality reduction is useful in 3D Gender classification with our DSF-based de-

descriptions except the *3D-grad.*, even with the expression changes in the 3D faces.

Table 4.16 – *confusion matrix of 3D-avg.*

	Asian	Non-Asian
Asian	77.90%	22.10%
Non-Asian	1.29%	98.71%
<i>Recognition rate =92.40%</i>		

Table 4.17 – *confusion matrix of 3D-sym.*

	Asian	Non-Asian
Asian	85.40%	14.60%
Non-Asian	2.91%	97.09%
<i>Recognition rate =93.55%</i>		

Table 4.18 – *confusion matrix of 3D-grad.*

	Asian	Non-Asian
Asian	53.17%	46.83%
Non-Asian	10.64%	89.36%
<i>Recognition rate =78.40%</i>		

Table 4.19 – *confusion matrix of 3D-spat.*

	Asian	Non-Asian
Asian	66.44%	33.56%
Non-Asian	2.12%	97.88%
<i>Recognition rate =88.36%</i>		

For Age estimation, as shown in Fig. 4.7 (C), we achieve < 4.6 years MAE for each description, when 5 principal components are used. With the *3D-avg.* description, we achieve 3.98 years MAE with 45 principal components. With the *3D-sym.* description, we achieve 3.99 years MAE with 15 principal components. With the *3D-grad.* description, we achieve 4.48 years MAE with 35 principal components. With the *3D-spat.* description, we achieve 4.23 years MAE with 25 principal components. For the results in each age group, as shown in Table 4.20, the results are significantly better in the young age groups, than in the old age groups. In terms of cumulative scores, as shown in Fig. 4.8, we achieve about 73.1 – 79.1% acceptance with a 5 years estimation error, and 91.3 – 92.5% acceptance with a 10 years error.

Table 4.20 – *Expression-Independent age estimation details with PCA transformed features*

	3D-avg.	3D-sym.	3D-grad.	3D-spat.	# subjects
≤ 20	3.43	3.49	4.08	3.85	185
(20,30]	2.28	2.35	2.68	2.43	246
(30,40]	8.85	8.03	9.64	8.48	20
≥ 40	23.42	23.62	23.60	24.64	15
All	3.98	3.99	4.48	4.13	466

4.4.2 CFS-based Expression-Independent Experiments

In parallel with the PCA method, we test also the CFS method for feature dimensionality reduction in the related recognition tasks under the

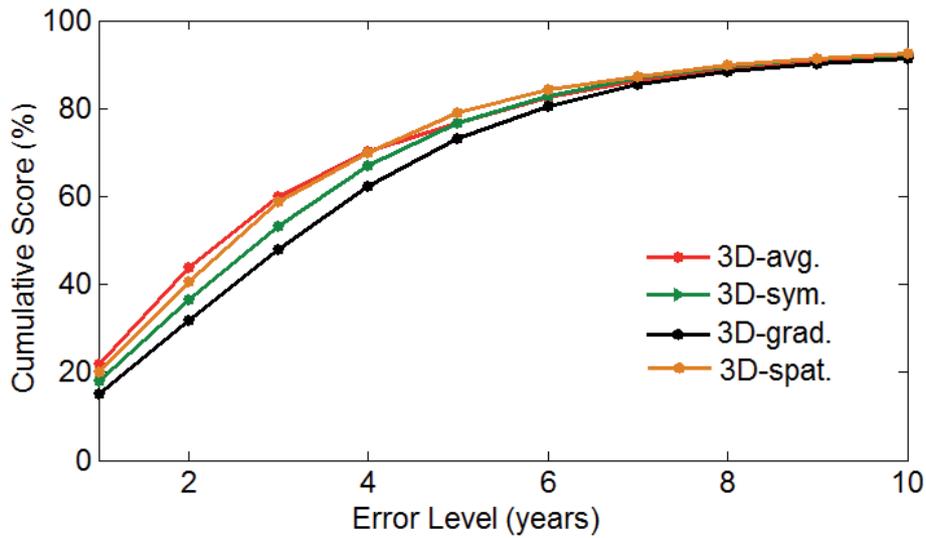


Figure 4.8 – Cumulative Score of age estimation with PCA transformed features

Expression-Independent setting. Rather than performing CFS based feature selection directly with the face descriptions of all the 4007 scans, we use the feature selection results on the 466 earliest scans to achieve feature dimensionality reduction. That means, in the Expression-Independent setting, the same subset of features are selected for each face description and attribute, than in the Expression-Dependent setting. With this, we have avoided the computational difficulty in feature selection with very big size data, and also got the opportunity of testing the generalization ability of the selected features with facial expression changes. After feature selection, we perform LOPO facial soft-biometric recognition experiments with Random Forest. The results are shown in Fig. 4.9.

In Fig. 4.9, for comparison purpose, the recognition results with the transformed PCA features on the 4007 scans are also shown in each sub-figure. These results are marked in blue color and labeled as *DSF-PCA* in the legends. For Gender classification, as shown in Fig. 4.9 (A), the performance with selected features (*DSF-SEL*) is 1-3% higher than that with the PCA transformed features. The Gender classification rate is always above 80% with each descriptor. With the selected features of *3D-avg.* and *3D-sym.*, we achieve 91.46% and 90.61%, respectively. With the *3D-grad.* description, we achieve 87.34% classification rate. For the *3D-spat.* description, the classification rate is 84.87%. In Ethnicity classification, as shown

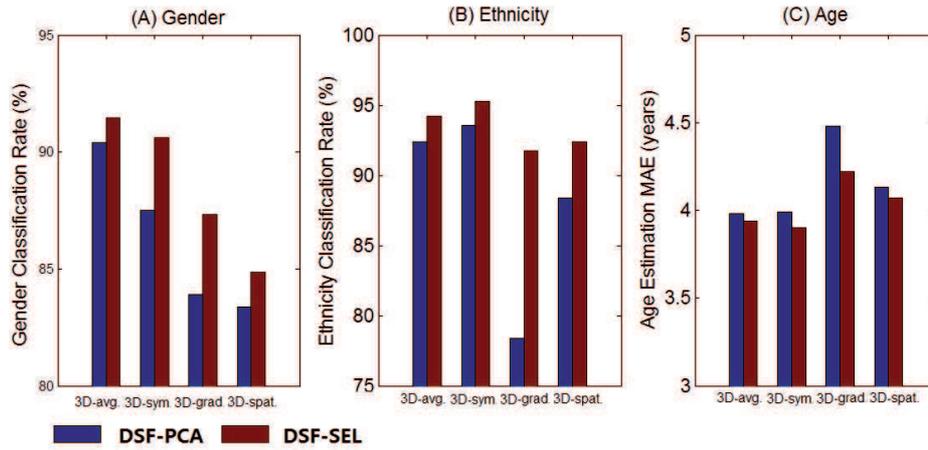


Figure 4.9 – Expression-Independent Experiments with CSF selected features

in Fig. 4.9 (B), the performance with selected features (*DSF-SEL*) is always $> 90\%$ and 2-10% higher than that with the PCA transformed features. With the selected features of *3D-avg.* and *3D-sym.*, we achieve 94.21% and 95.26%, respectively. With the *3D-grad.* description, we achieve 91.71% ethnicity classification rate. For the *3D-spat.* description, the ethnicity classification rate is 92.38%. For Age estimation, the results shown in Fig. 4.9 (C) also demonstrate clear enhancement with feature selection, than with the PCA transformed features. The MAEs are always lower compared with the transformed PCA features. We achieve 3.94 and 3.90 years MAE with the *3D-avg.* and *3D-sym.* descriptions. With the *3D-grad.* description, we achieve 4.22 years MAE. For the *3D-spat.* description, the MAE rate is 4.07. When looking at the MAEs in each age group, as shown in Table 4.21, the MAEs in the (30,40] age group are significant lower than the results with PCA. In terms of cumulative score, as shown in Fig. 4.10, the CFS selected features achieve 76-78% acceptance with a 5 years estimation error, and 93% acceptance with a 10 years estimation error. In comparison, the cumulative scores of age estimation with the CSF selected features are slightly better than the CSs from the PCA transformed features.

4.4.3 Summary

In this section, we have explored the performance of the DSF features under the Expression-Independent setting, in combination of two feature

Table 4.21 – Expression-Independent age estimation details with CFS selected features

	3D-avg.	3D-sym.	3D-grad.	3D-spat.	# subjects
≤ 20	4.10	4.03	4.37	4.18	185
(20,30]	2.40	2.44	2.73	2.61	246
(30,40]	7.25	6.56	7.25	7.10	20
≥ 40	23.93	23.78	23.89	23.38	15
All	3.94	3.90	4.22	4.07	466

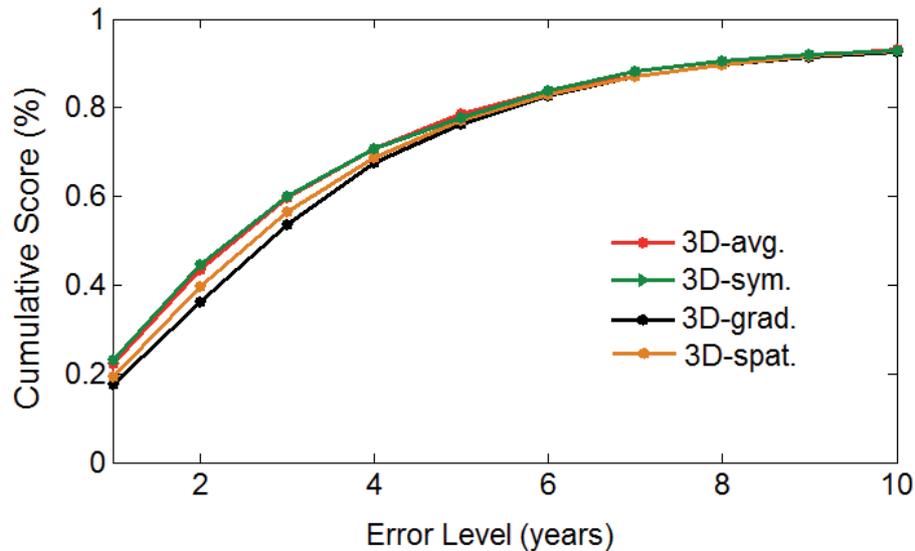


Figure 4.10 – Cumulative Score of age estimation with CFS selected features

dimensionality reduction methods. With the PCA method, in Expression-Independent LOPO cross validation with different DSF description, we achieve always $> 80\%$ gender classification rate, $> 88\%$ ethnicity classification rate (except *3D-grad.*), and < 4.5 years MAE. With the CFS method applied on each DSF description, we achieve always $> 85\%$ gender classification rate, $> 90\%$ ethnicity classification rate, and < 4.2 years MAE. These results are comparable with the results from Expression-Dependent experiments with the original high dimensional DSF features, PCA transformed features and the CFS selected features. It demonstrates that the proposed facial soft-biometric recognition approaches with the DSF features are effective and robust to facial expression changes. The results also shows that in both the Expression-Dependent and Expression-Independent settings, the CFS always works better than the PCA method for the recognition of these facial soft-biometrics. In later experimental

analysis, we keep only the CFS method for feature dimensionality reduction.

4.5 FEATURE LEVEL FUSION EXPERIMENTS

As explained before, with the four types of DSF features derived from high-level facial morphology cues, we are interested in the *early fusion* (or *feature-level fusion*) method for discovering that whether the four descriptions are complimentary or not in the facial soft-biometric recognition tasks. Compared with the *late fusion* method, the *early fusion* serves better our goal, as it contains completely the information from each description, and do not need separate learning stage. In our work, we establish the *early fusion* by concatenating our four types of DSF features. Under the Expression-Dependent setting, we experiment with the original DSF features (DSF466) and the selected DSF features (SEL466). For the Expression-Independent setting, we explore the fusion method with the selected DSF features (SEL4007). The corresponding recognition results of these soft-biometrics are shown in Fig. 4.11. For comparison, in each subplot, the highest result achieved by individual DSF description (No-fusion) is shown as blue bar, on the left of the fusion result.

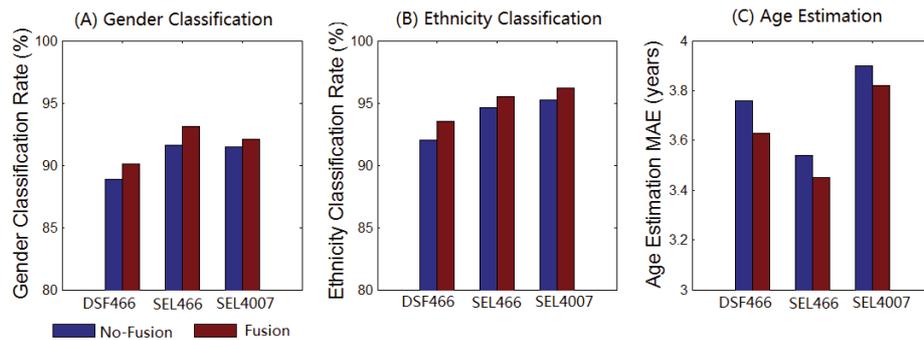


Figure 4.11 – Feature level fusion results for facial soft-biometric recognition

Table 4.22 – confusion matrix of ED (Gender)

	female	male
female	91.63%	8.37%
male	5.70%	94.30%
Recognition rate =93.13%		

Table 4.23 – confusion matrix of EI (Gender)

	female	male
female	92.80%	7.20%
male	8.48%	91.52%
Recognition rate =92.11%		

Table 4.24 – *confusion matrix of ED (Ethnicity)*

	Asian	Non-Asian
Asian	85.71%	14.29%
Non-Asian	1.41%	98.59%
<i>Recognition rate =95.49%</i>		

Table 4.25 – *confusion matrix of EI (Ethnicity)*

	Asian	Non-Asian
Asian	90.11%	9.89%
Non-Asian	1.11%	98.89%
<i>Recognition rate =96.23%</i>		

4.5.1 Expression-Dependent Experiments

As shown in Fig. 4.11 (A), the Gender classification result improves from 88.84% to 90.12% with the fusion of the original DSF features. The fusion of the selected DSF features achieves 93.13% classification rate of Gender, which is 1.5% higher than the result achieved by the selected features of $3D$ -avg. These fusion result with the selected features is detailed in Table 4.22. We observe that the results for male and female groups are effective ($> 90\%$), and comparable to each other. For Ethnicity classification, as shown in Fig. 4.11 (B), the results with the fusion of original DSF features achieves 95.49% classification rate. It outperforms the highest result achieved by single $3D$ -avg description by 1.54%. With the fusion of selected DSF features, we achieve 95.49% classification rate. This result outperforms the result of single $3D$ -avg description by 0.85%. In Table 4.24, the detailed results for Asian and Non-Asian groups are shown, considering the fusion of the selected features. The recognition results for both Asian and Non-Asian are effective ($> 85\%$), while the result for Non-Asian faces are still much higher than the result for the Asian faces. For Age estimation, as shown in Fig. 4.11 (C), the MAEs with fusion are always lower than the MAEs with each individual description. The MAE reduces to 3.63 years with the fusion of the original DSF features. When using the fusion of selected features, the MAE further reduces to 3.45 years. The group-wised age estimation MAEs of this result are shown in Table 4.26, labeled as 'SEL466'. Generally, the age estimation accuracy improves, especially in the ≤ 40 age groups. In terms of cumulative score, as shown in Fig. 4.12, it achieves 83.48% acceptance with a 5 years age estimation error, and 94.64% with a 10 years age estimation error. These results outperform significantly the CSs with individual DSF descriptions, as shown in Fig. 4.6. In summary of the fusion results under the Expression-Dependent

setting, the fusion results for Gender, Ethnicity and Age always outperform the individual face description. It demonstrates that, when facial expressions are not involved, the four types of face descriptions are complementary to each other in Gender, Ethnicity and Age recognition. Also, it shows the proposed feature-level fusion is effective in the related recognition tasks.

Table 4.26 – Age estimation details with the fusion of selected features

	≤ 20	(20,30]	(30,40]	≥ 40	ALL
SEL466	3.99	2.11	6.48	24.50	3.31
SEL4007	3.90	2.34	6.69	23.83	3.82
# subjects	185	246	20	15	466

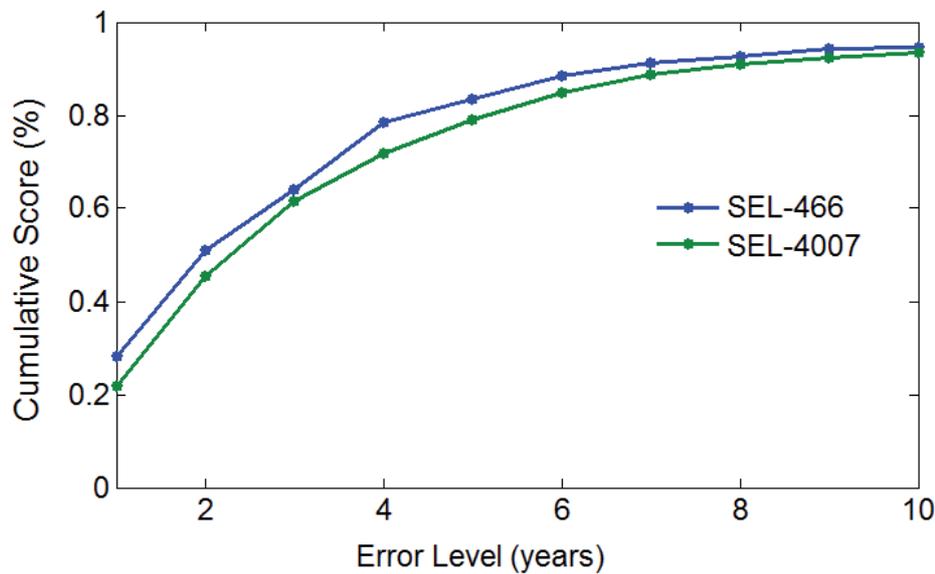


Figure 4.12 – Cumulative score of age estimation with fusion of selected features

4.5.2 Expression-Independent Experiments

Under the Expression-Dependent setting, the fusion method improves the recognition performance of facial soft-biometrics. While, with expression changes, is the fusion still helpful? With results labeled as *SEL4007* in each subplot of Fig. 4.11, we generate the positive answer. As shown in Fig. 4.11 (A), we achieve 92.11% Gender classification rate with the fusion of selected features, which is 0.65% higher than the result of using the single *3D-avg.* features. Compared with the fusion result in Expression-

Dependent setting, it has only a small decrease of 1%. It shows that our method is robust to expression changes in Gender classification. The confusion matrix of this result is shown in Table 4.23. It's clear that for both the male scans and the female scans, the recognition rates are effective ($> 91\%$), and very close to each other. For Ethnicity, as shown in Fig. 4.11 (B), we achieve 96.23% classification rate in the fusion. This result is 1% higher than with the single *3D-sym.* description, and it is also even 0.74% higher than the corresponding fusion result in the Expression-Dependent setting. It means that our method is also robust to expression changes in Ethnicity classification. In Table 4.25, the confusion matrix shows that for both the Asian and Non-Asian scans, the recognition rates are effective ($> 90\%$). The recognition rate for Non-Asian scans is 8% higher than the result for the Asian faces. While compared to the results in the Expression-Dependent experiments, as shown in Table 4.24, the difference of recognition rate between the Asian and the Non-Asian faces is reduced largely. For Age, as shown in Fig. 4.11 (C), we achieve 3.82 years MAE with the fusion of selected features. It is lower than the result with the single *3D-sym.* description. Although it still has a big gap to the 3.45 years MAE achieved in the Expression-Dependent setting, considering that the size of the dataset, it still demonstrates the robustness of our method to expression changes in Age estimation, to some extent. The concerning age estimation MAEs in each age group are shown in Table 4.26, labeled as 'SEL4007'. Again, in general, the age estimation accuracy improves in each age group, especially in the ≤ 40 age groups. In terms of cumulative score, as shown in Fig. 4.12, it achieves 79.03% acceptance with a 5 years age estimation error, and 93.53% with a 10 years age estimation error. These results outperform significantly the CSs with individual DSF descriptions, as shown in Fig. 4.10. In summary of the fusion results under the Expression-Independent setting, the fusion results for Gender, Ethnicity and Age always outperform the corresponding individual face description. It demonstrates that, even with the facial expressions, the four types of face descriptions are complimentary to each other in Gender, Ethnicity and Age recognition. Also, it confirms that the proposed

feature-level fusion is effective in the recognition tasks of the concerning soft-biometrics.

4.5.3 Summary

From the experiments above, we find that under both the Expression-Dependent and the Expression-Independent settings, the facial soft-biometric recognition performances are always higher than those results with single DSF descriptions. Thus, we conclude that, the feature level fusion of our DSF descriptions improves the performance of facial soft-biometric recognition. It means that our DSF descriptions are complimentary in discriminating these facial soft-biometrics. Furtherly, it means that the underlying morphology cues, namely the face Averageness, the face Symmetry, the Spatial configuration and the local gradient, are complimentary in conveying the related cues of Gender, Ethnicity and Age.

4.6 ROBUSTNESS TO FACIAL VARIANTS

The robustness to various facial variants is an important issue of facial soft-biometric recognition approaches. It measures the robustness of the recognition performance when imposing the challenge of other facial variants. Thus, following the above experiments and discussions, a natural question is, how much do the facial variants influence our facial soft-biometric recognition performances? With the FRGCv2 dataset, we have the opportunity to answer this question, as it includes various facial variants, the gender, the ethnicity, the age, and the facial expressions. In the following, we further analyze the feature-level fusion results under the Expression-Independent setting, to evaluate the robustness of our approaches to various facial variants. For Gender, we consider two groups, Male (M) and Female (F). For Ethnicity, we consider the Asian group (AS) and Non-Asian group (NS). For age, we make six age groups from age 10 to 70 with equal interval. For Expression, we take the six expression groups of FRGCv2, namely the 'Blankstare' (BL), 'Happiness' (HP), 'Surprise' (SP), 'Disgust' (DG), 'Sadness' (SD), and the rest labeled as 'Other'

(OT). The details of these facial variants in FRGCv2 dataset are depicted in Fig. 4.27-4.30.

Table 4.27 – Expression groups in FRGCv2

Expression	BL	HP	SP	DG	SD	OT
# scans	2364	378	202	342	178	541

Table 4.28 – Gender groups in FRGCv2

Gender	female	male
# scans	1847	2158

Table 4.29 – Ethnicity groups in FRGCv2

Ethnicity	Asian	Non-Asian
# scans	1213	2792

Table 4.30 – Expression groups in FRGCv2

Age	(10,20]	(20,30]	(30,40]	(40,50]	(50,60]	(60,70]
# scans	1361	2222	253	86	75	8

In Gender classification, Ethnicity, Age and facial Expressions are considered as influencing facial variants. In Fig. 4.13 (A),(B),and (C), we present the Gender recognition results in different Ethnicity, Facial Expression and Age groups, respectively. In each subplot of Fig. 4.13, the blue bars indicate the Gender classification rate, while the red lines show the number of scans in the concerning group. In all the subgroups concerning all the three facial variants, the Gender classification rates are almost always around 90%. Only in the age groups of (50,60] and (60,70], the Gender classification reach a lower but still effective performance of about 85%. The Gender classification performances do not change a lot in the subgroups of each variant, although the number of scans in these subgroups changes vibrantly. Thus, in conclusion, our Gender classification approach has good robustness to Ethnicity, Expression and Age changes in faces.

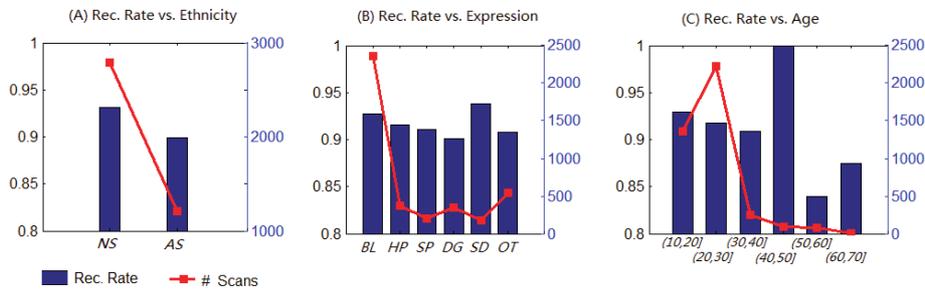


Figure 4.13 – Gender recognition rate concerning facial Variants

In Fig. 4.14 (A),(B),and (C), we present the Ethnicity recognition results

in different Gender, Facial Expression and Age groups, respectively. From Fig. 4.14, we observe that, in all the subgroups concerning all the three facial variants, the Ethnicity classification rates are always around 95%. Despite the big changes of the number of scans for the subgroups of each variant, the performance of Ethnicity classification stays relatively stable. Thus, we conclude that our Ethnicity classification approach has good robustness to Gender, Expression and Age changes in faces.

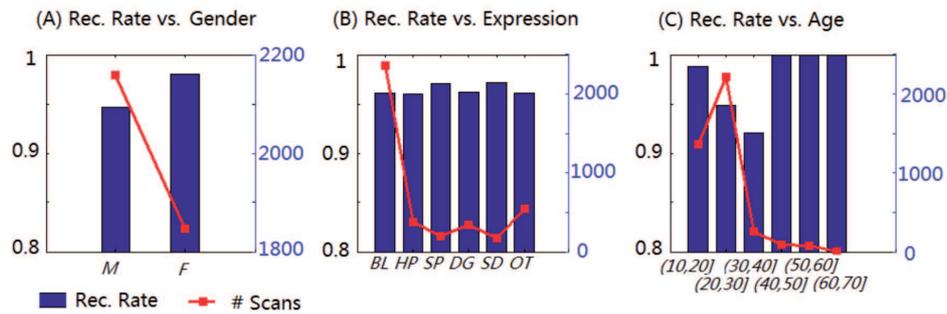


Figure 4.14 – Ethnicity recognition rate concerning facial Variants

In Fig. 4.15 (A),(B),and (C), we present the Age estimation MAEs in different Gender, Ethnicity and Facial Expression groups, respectively. From Fig. 4.15, we observe that, in all the subgroups concerning all the three facial variants, the Age estimation MAEs are always below 4.5 years. Despite the big changes of the number of scans for the subgroups of each variant, the performance of Age estimation stays relatively effective. While, unlike for Gender and Ethnicity, the Age estimation performance changes more vibrantly in different groups of a concerning facial variant. In conclusion, we conclude that our Age estimation approach stays effective with Gender, Expression and Age changes in faces, while the performance also suffers from these variants, to some extend.

In conclusion of the analysis between recognition performances and the influential variants, our facial soft-biometric recognition approaches are relatively robust to the facial variants, including Gender, Ethnicity, Age and facial Expression.

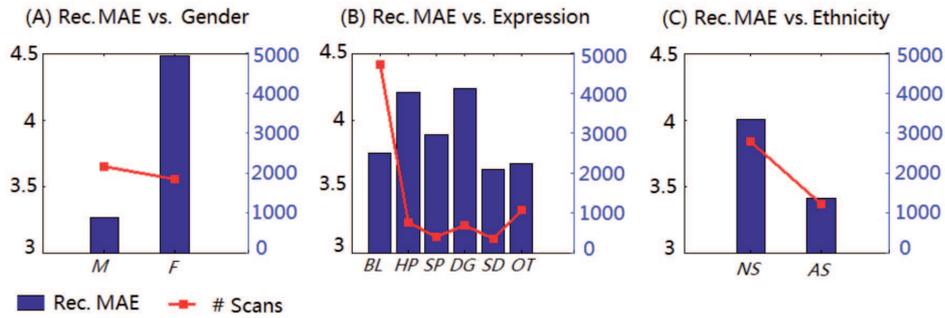


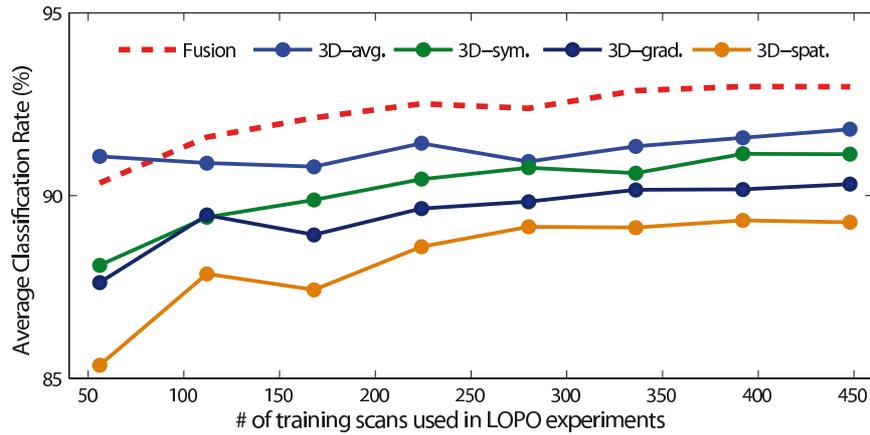
Figure 4.15 – Age recognition MAE concerning facial Variants

4.7 ROBUSTNESS TO TRAINING SIZE

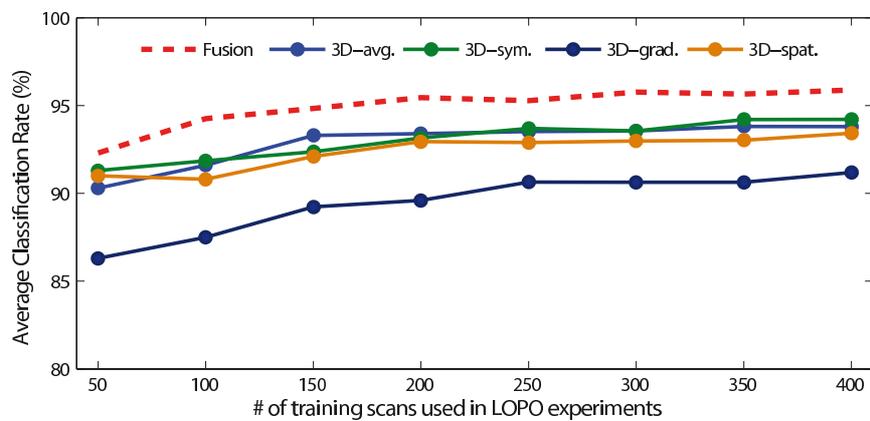
Beyond the recognition accuracy, another important issue of these algorithms is the dependency to the size of training set. That is to say, at least how many training instances are needed to obtain a good performance, and how dose the size of training data influence the experimental performance? Following this perspective, we perform a set of recognition experiments by varying the size of the training set, for each facial soft-biometric. These experiments are performed with the selected DSF features of the first 466 scans of FRGCv2. Each time, a specific number of training instances are randomly picked out from the DSFs from the 466 scans, and LOPO experiments are performed within this subset². To generate a statistical view of the performance, for each specified number, we perform 20 times of the experiment. Then we use the average accuracy to evaluate the overall performance. The corresponding recognition performances for Gender, Ethnicity and Age are shown in Figure 4.16a, Figure 4.16b and Figure 4.16c, respectively.

The results in Figure 4.16a, Figure 4.16b and Figure 4.16c show that, for all the three attributes, the recognition accuracy increases when increasing the number of training scans. It demonstrates that the increase of training scans gives benefits to the recognition performance. Also, even with very few training data, the recognition performances are still effective.

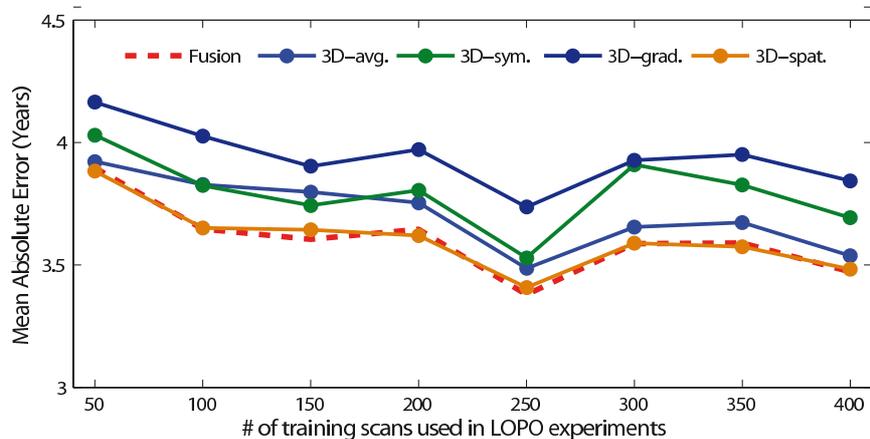
²In Gender classification, the number of scans used in the LOPO experiments ranges from 56 to 448 (8 equally interpolated integers), and always in respect of the Asian/Non-Asian scans ratio as 1:3. For Ethnicity classification, the number of scans used in the LOPO experiments ranges from 50 to 400 (8 equally interpolated integers), in respect of the Male/Female scans ratio as 1:1. In Age estimation, the number of scans used in the LOPO experiments ranges from 50 to 400 (8 equally interpolated integers).



(a) Gender Classification



(b) Ethnicity Classification



(c) Age Estimation

Figure 4.16 – Recognition accuracy when varying the # of training scans

In gender classification as shown in Figure 4.16a, > 85% correct classification rate is achieved with only 56 scans for all the descriptions. With the **3D-avg.** description, it even achieves > 90%. For ethnicity classification as shown in Figure 4.16b, all the descriptions give > 86% correctness

with only 50 scans. Except only for the **3D-spat.**, all other descriptions achieve $> 90\%$ correctness. For age estimation as shown in Figure 4.16c, we achieve $< 3.9\%$ years MAE with 50 scans in the LOPO experiment. These results show that our approach works well even with very limited training data in these recognition tasks. Also, these figures confirm again that the feature-level fusion outperforms each single description, for all the concerning facial soft-biometrics here.

4.8 COMPARISON WITH THE STATE-OF-THE-ART

In the literature of 3D-based facial soft-biometric recognition, no work has been issued considering the correlations of Gender, Ethnicity and age. There are several works which perform 3D-based Gender and Ethnicity classification, as described in Table 4.31. Up to now, there is no literature in 3D-based age estimation, except our work in [97], which works on 466 earliest scans of FRGCv2. From Table 4.31, we find that most of the related works in the state-of-the-art use the 10-fold cross-validation. For better comparison with the related works, we also present in Table 4.31 the Gender and Ethnicity recognition results under the 10-fold cross-validation experiments³. The results are provided in the form of the mean classification rate plus the standard deviation of the classification rates in the 10 folds of experiments.

As shown in Tab.4.31, for Gender classification, the works closely related to ours are presented in [9, 91, 36, 93, 52], which are also tested on the FRGCv2 dataset. Under the Expression-Dependent setting, our Gender classification rate on the 466 earliest scans of FRGCv2 (95.06%) outperforms significantly the results of *Ballihi et al.* (86.05%) presented in [9]. Our result is about 2% lower than *Gilani et al.* in [36] (97.05%). With all the 4007 scans of FRGCv2, we achieve 93.11% Gender classification rate. This results is comparable to the results of *Toderici et al.* in [98] (93.5%), *Wang et al.* in [93] (93.7%), and is about 3% lower than the result of *Gilani et al.* in [36] (96.12%). Thus, compared with the literature, the only work

³We note that this is the only case that we use 10-fold cross-validation in this present thesis. In other parts of the thesis, all the experiments follow the LOPO protocol.

Table 4.31 – comparison with state of the art

Gender Classification (Male / Female)							
<i>Author</i>	<i>Dataset</i>	<i>Auto</i>	<i>Features</i>	<i>Classifiers</i>	<i>Setting</i>	<i>Results</i>	<i>Modality</i>
<i>Ballihi et al. [9]</i>	466 scans of FRGCv2	Yes	Facial curves	Adaboost	10-fold	86.05%	shape
<i>Gilani et al. [36]</i>	466 scans of FRGCv2	Yes	landmark distances	LDA classifier	10-fold	97.05%	shape
	4007 scans of FRGCv2	Yes	landmark distances	LDA classifier	10-fold	96.12%	shape
<i>Toderici et al. [91]</i>	4007 scans of FRGCv2	Yes	Wavelets	Polynomial-SVM	10-fold	93.50% ±0.045	shape
<i>Wang et al. [93]</i>	4007 scans of FRGCv2	No	3D coordinates	RBF-SVM	5-fold	93.70% ±0.02	shape + texture
<i>Huang et al. [52]</i>	3676 scans of FRGCv2	No	LCP features	Adaboost	10-fold	95.50% ±0.03	shape + texture
Our work	466 scans of FRGCv2	Yes	DSF features	Random Forest	10-fold	95.06% ±0.027	shape
	4007 scans of FRGCv2	Yes	DSF features	Random Forest	10-fold	93.11% ±0.035	shape
Ethnicity Classification (Asian / Non-Asian)							
<i>Author</i>	<i>Dataset</i>	<i>Auto</i>	<i>Features</i>	<i>Classifiers</i>	<i>Setting</i>	<i>Results</i>	<i>Modality</i>
<i>Zhong et al. [104]</i>	4007 scans of FRGCv2	No	LVC features	membership probability	–	82.38%	shape
<i>Toderici et al. [91]</i>	3676 scans of FRGCv2	Yes	Wavelets	Polynomial-SVM	10-fold	99.50% ±0.01	shape
<i>Huang et al. [52]</i>	3676 scans of FRGCv2	No	LCP features	Adaboost	10-fold	99.60% ±0.01	shape + texture
Our work	466 scans of FRGCv2	Yes	DSF features	Random Forest	10-fold	96.78% ±0.023	shape
	4007 scans of FRGCv2	Yes	DSF features	Random Forest	10-fold	96.45% ±0.033	shape

that outperforms our approach in result is presented by *Gilani et al.* in [36]. While, their work depends on the accurately detected facial landmarks for feature extraction. In comparison, our proposed method is independent of facial landmarks.

For Ethnicity classification, the nearest works with ours are presented by *Zhong et al.* in [104], *Toderici et al.* in [91], and *Huang et al.* in [52]. With the 466 earliest scans of FRGCv2, we achieve 96.56% Ethnicity classification rate in 10-fold cross-validation experiments. This is the first result in the literature with the 466 earliest scans of FRGCv2. With all the 4007 scans of FRGCv2, we achieve 96.45% Ethnicity classification rate. This result is much higher than the results of *Zhong et al.* in [104] (82.38%). Compared with *Toderici et al.* in [91] and *Huang et al.* in [52], our Ethnicity classification rate is 3% lower. However, their results are based on the 3676

Asian and White scans of FRGCv2. The scans of the remaining subjects are not considered, which correspond to the 28 scans of 6 Black-or-African American subjects, the 113 scans of 13 Hispanic subjects, and the 97 scans of 16 subjects whose Ethnicity is unknown. In contrast, we consider all the 4007 scans of FRGCv2 which cover all the provided Ethnicity types. Thus, compared to the work of *Toderici et al.*, we encounter significantly more complicated Ethnicity challenges.

4.9 CONCLUSION

In conclusion of this chapter, we have the following achievements. *First*, we have proposed to use four types of facial morphology cues, concerning the face Averageness, the Symmetry, the Spatial configuration and the local shape Gradient, for Gender, Ethnicity and Age recognition. These cues are closely related to these facial soft-biometrics in 3D facial surfaces. *Secondly*, with the facial morphology cues, we have proposed to extract four types of geometric descriptions from facial surfaces, through Riemannian shape analysis of facial curves on a specific Riemannian manifold. The extracted features capture densely the facial morphology cues on the 3D surface. *Thirdly*, with the extracted geometric features, we have developed effective facial soft-biometrics recognition approaches, which are also robust to the facial expression changes. In the LOPO Expression-Dependent experiments, we have achieved 95.05% Gender classification rate, 96.78% Ethnicity classification rate, and 3.31 years MAE for age estimation. In the Expression-Independent experiments, the recognition accuracy stays comparable. We have achieved 93.11% Gender classification rate, 96.45% Ethnicity classification rate, and 3.82 years MAE in age estimation. The results from both the Expression-Dependent and -Independent settings have demonstrated that the 3D shape of faces can reveal the Gender, Ethnicity and Age of subjects, and the facial cues related to Gender, Ethnicity and Age exist in terms of the face Averageness, the Symmetry, the Spatial configuration and the local shape Gradient. *Fourthly*, we have proposed a feature-level fusion method, with which we have effectively enhanced the recognition results of all the concerning soft-biometrics. It demon-

strates that the proposed features are complimentary to each other in the related recognition tasks. Furtherly, it means that the face Averageness, the Symmetry, the Spatial configuration and the local shape Gradient captures complimentary information of Gender, Ethnicity and Age in the 3D faces. Last, we have demonstrated that the proposed facial soft-biometric recognition algorithms are robust to the facial variances, and the size of training set.

JOINT FACIAL SOFT-BIOMETRICS RECOGNITION

SOMMAIRE

5.1	INTRODUCTION	82
5.2	CORRELATED FACIAL ATTRIBUTES RECOGNITION	83
5.2.1	Gender Classification Experiments	83
5.2.2	Ethnicity Classification Experiments	85
5.2.3	Age Estimation Experiments	86
5.3	HOW MUCH ARE GENDER, ETHNICITY AND AGE CORRELATED ?	89
5.4	HOW TO BENEFIT FROM THEIR CORRELATION IN REAL-WORLD LIKE APPLICATIONS ?	92
5.5	CONCLUSION	95

5.1 INTRODUCTION

In the previous chapter, we have made comprehensive studies of facial soft-biometric recognition using the proposed geometric features. In these studies, each facial soft-biometric is recognized individually without giving consideration of the others. However, as stated before, these soft-biometrics do not only demonstrate clear cues in shape for recognizing themselves, but also co-exist naturally in the face. According to facial anthropometry study [106], Gender, Ethnicity and Age interact with each other in characterizing the face shape. In public perception, age perception is in correlation with gender [2] and ethnicity [102]. In facial demographic classification, the correlation between gender and ethnicity has also been noticed [37, 33, 37].

In the literature of 2D based facial soft-biometric recognition, several works have been done concerning the correlations among these soft-biometrics, for ethnicity-specific gender classification [33, 28], gender-specific ethnicity classification [28], gender-specific age estimation [81, 38, 64, 61, 43], ethnicity-specific age estimation [43], ethnicity&gender-specific age estimation [43]. No work has given thorough exploration to the correlations of all the three soft-biometrics. Further, in the 3D domain, there is still no work giving consideration to the correlations of these soft-biometrics. Considering this, in this chapter, we propose to give a thorough exploration of the correlations of these soft-biometrics in their recognition tasks.

To give consideration to their correlations in the experiments, we define the *biometric-specific* experimental settings as following. For *Gender-specific* setting, the 466 subjects are separated into Male group (263 subjects) and Female group (203 subjects) first, and then we experiment on each group separately. For *Ethnicity-specific* setting, we separate the 466 scans into Asian group (112 subjects, correspond to the Subjects labeled as Asian, Asian-southern and Asian and Middle-east) and Non-Asian group (the rest 354 subjects) first, and then do experiments on each ethnicity group separately. For *Age-specific* setting, we separate the 466 subjects into older group (≥ 26 years, 107 subjects) and younger group (≤ 25 years,

359 subjects) first, and then perform LOPO experiments on the younger and older groups, separately. By working on a specific population group, the relationship between the studied soft-biometric and the population groups which is defined using another facial attribute, is able to be revealed. These groups partitions are always valid in this section, except with clear declaration of changes.

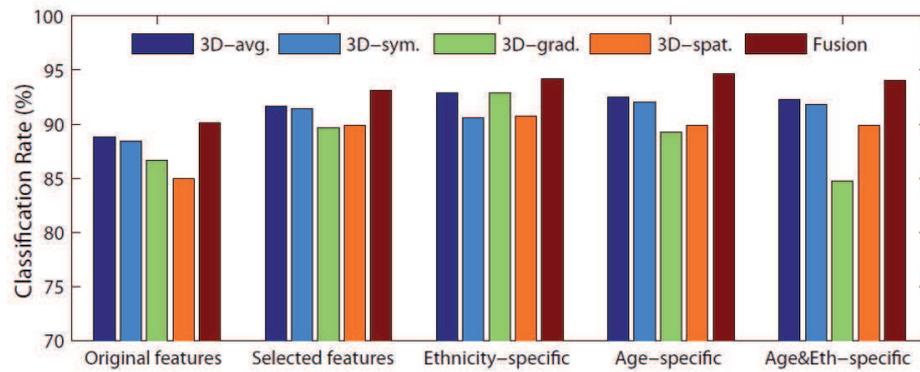
5.2 CORRELATED FACIAL ATTRIBUTES RECOGNITION

For the correlated facial soft-biometric recognition, we use also the *Leave-One-Person-Out (LOPO)* cross-validation protocol. In the *biometric-specific* experiments, the CFS selected features for each DSF description, and their fusion are employed as features. The experimental analysis for Gender, Ethnicity and Age are shown in the following.

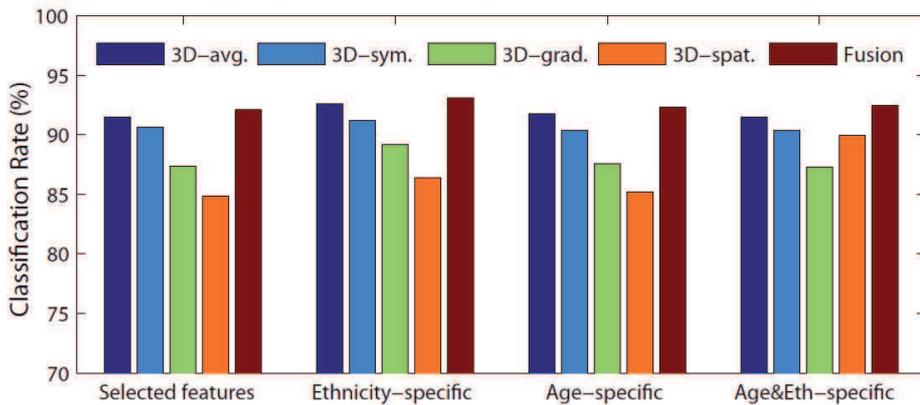
5.2.1 Gender Classification Experiments

To give consideration of Ethnicity and Age information in Gender classification, we perform the *Ethnicity-specific* and *Age-specific* gender classification experiments with the selected features of each DSF description. We have also done the experiments considering both ethnicity and age. In the *Eth.&Age-specific* setting, we perform LOPO gender classification experiments with the scans from the same ethnicity and age groups. The Gender classification results in both the Expression-Dependent and Expression-Independent experiments are shown as bar-plots in Fig. 5.1. In each panel of Fig. 5.1, the y-axis shows the classification rate in LOPO experiment, and the x-axis shows the different experimental settings. For comparison, the corresponding experimental results with the selected features in the LOPO experiments are also shown in each panel.

As shown in Fig. 5.1 (a), in the Expression-Dependent experiments, when considering the ethnicity information in the *Ethnicity-specific* setting, the gender classification results are improved in general than with the selected features. **It shows that Asian and Non-Asian people show different gender patterns.** When considering the age information in the *Age-specific* setting, a stronger improvement is shown. **It indicates that**



(a) Expression-dependant gender classification results



(b) Expression-independant gender classification results

Figure 5.1 – *Gender classification results under Expression-dependent and Expression-Independent settings.* Features, 3D-avg.: Averageness — 3D-sym.: Bilateral Symmetry — 3D-grad.: Gradient — 3D-spat.: Spatial — Fusion: their fusion by concatenation. Features processing, Original features: No feature selection applied — Selected features: Correlation-based features selection applied before classification. Settings, Ethnicity-specific: Selected features within each ethnicity group — Age-specific: Selected features within each age group — Age&Eth-specific: Selected features within the same ethnicity and age group.

people of different age have different gender patterns. When considering both *Ethnicity-specific* and *Age-specific*, termed *Age&Eth-specific*, the accuracy is generally higher than the selected features and quite comparable to *Ethnicity-specific* and *Age-specific*. In addition, the fusion of the features always outperforms individual features in all the settings. The highest gender classification rate, 94.64% and 94.21%, are achieved by the *Fusion* under the *Age-specific* and *Ethnicity-specific* settings, respectively. These findings are furtherly confirmed by in the Expression-Independent experiments, as shown in Fig.5.1 (B). For each description, the gender classification performance is always higher when considering Ethnicity and

Age information. The fusion of these features always outperforms each individual description, and achieves an 93.13% accuracy in the *Ethnicity-specific* setting. These results show also that the expressions variations affect slightly the method performance.

5.2.2 Ethnicity Classification Experiments

In Ethnicity classification, we perform the *Gender-specific* and *Age-specific* ethnicity classification experiments to give consideration of Gender and Age information. We have also done the experiments considering both gender and age. In the *Gen&Age-specific* setting, we perform LOPO ethnicity classification experiments with the scans from the same gender and age groups. The ethnicity classification results in both the Expression-Dependent and Expression-Independent experiments are shown as barplots in Fig. 5.2. Similar to the previous gender classification, for comparison usage, the corresponding experimental results with the selected features in the LOPO ethnicity classification experiments are also shown in each panel of the Fig. 5.2.

As shown 5.2 (A), in the Expression-Dependent Experiments, under the *Gender-specific* and *Age-specific* settings, the ethnicity classification results are slightly higher than the results using directly the selected features. **It shows that the Male and the Female have different ethnicity patterns, and people of different age have different ethnicity patterns.** Here, the highest ethnicity classification rates of 95.71% and 95.49% are achieved by the *3D-avg.* description and the fusion, respectively. Again, like in gender classification, the fusion of these features generally outperforms each individual description in each experimental setting. These results are confirmed in Fig. 5.2 (B) with a higher accuracy of 96.6%. It demonstrates the robustness of the proposed approach against the facial expressions in ethnicity classification. Roughly speaking, according to the results presented above, Ethnicity (Asian and Non-Asian) classification is influenced by gender and age factors.

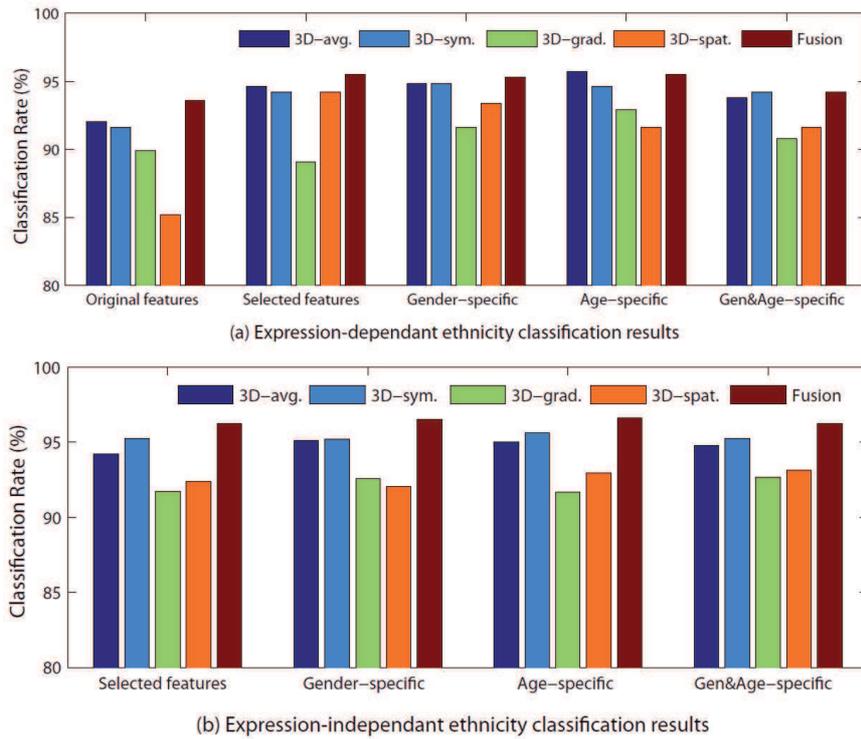
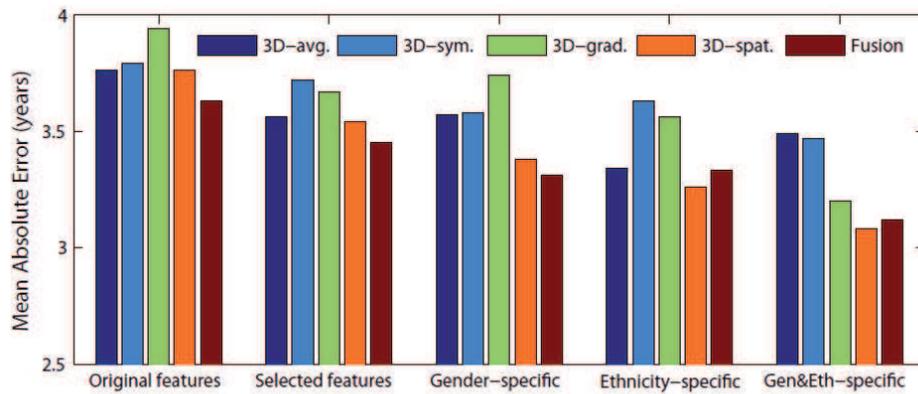


Figure 5.2 – *Ethnicity classification results under Expression-dependent and Expression-Independent settings.* Features, 3D-avg.: Averageness — 3D-sym.: Bilateral Symmetry — 3D-grad.: Gradient — 3D-spat.: Spatial — Fusion: their fusion by concatenation. Features processing, Original features: No feature selection applied — Selected features: Correlation-based features selection applied before classification. Settings, Gender-specific: Selected features within each gender group — Age-specific: Selected features within each age group — Age&Gen-specific: Selected features within the same gender and age group.

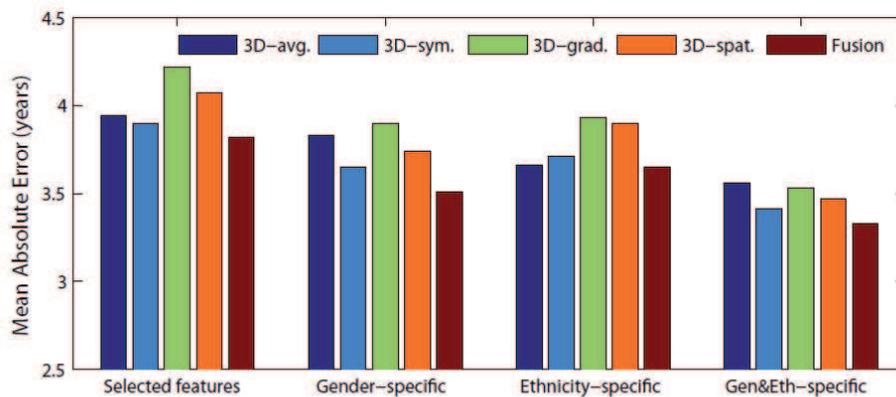
5.2.3 Age Estimation Experiments

In the following, we perform *Gender-specific* and *Ethnicity-specific* experiments to give consideration of the gender and ethnicity information in age estimation. Similar to the previous experiments for Gender and Ethnicity classifications, here again, we compare the results achieved with the results from the selected features, and results reported under *Gender-specific* and *Ethnicity-specific* settings. In the *Gen&Eth-specific* setting, we perform LOPO experiments with the scans from the same gender and ethnicity groups, to give consideration of both the gender and ethnicity information in age estimation. The experimental results are shown as MAEs in Fig. 5.3.

As shown in Fig. 5.3 (A), in the Expression-Dependent experiments,



(a) Expression-dependant age estimation results (MAE)



(b) Expression-independent age estimation results (MAE)

Figure 5.3 – Age estimation accuracy under Expression-dependent and Expression-independent settings. Features, 3D-avg.: Averageness — 3D-sym.: Bilateral Symmetry — 3D-grad.: Gradient — 3D-spat.: Spatial — Fusion: their fusion by concatenation. Features processing, Original features: No feature selection applied — Selected features: Correlation-based features selection applied before classification. Settings, Gender-specific: Selected features within each gender group — Ethnicity-specific: Selected features within each age group — Eth&Gen-specific: Selected features within the same gender and ethnicity group.

under the *Ethnicity-specific* and the *Gender-specific* settings, the MAEs are significantly lower than the MAEs from the selected features. **It means that the Asian and the Non-Asian people have different facial aging patterns, and also, Male and Female people are aging differently.** When considering both gender and ethnicity in the *Gen&Eth* setting, the MAEs are even lower than the ones under the *Ethnicity-specific* and *Gender-specific* settings. It demonstrates that **the combination of gender and ethnicity information gives the strongest improvement to age estimation performance.** Also, the **Fusion** generally provides the highest performances in each setting. The lowest MAEs are achieved in the *Gen&Eth-specific*

setting, by **spat.** and **Fusion** at 3.08 and 3.12 years, respectively. These observations are confirmed by the results of the Expression-Independent experiments, as depicted in Fig 5.3 (B). When using the gender and ethnicity information together, the improvement of performance is the most significant. The *Fusion* of the descriptions always outperform individual descriptions, and achieves the lowest MAE of 3.33 years in the *Gen&Eth-specific* setting. We note, despite the facial expression variations, which affect significantly the facial shape, our algorithm still provides a high accuracy.

Following the 2D-based literature, we show in Table 5.1 the age estimation accuracy in each age group. We find that, no matter in Expression-Dependent or Expression-Independent experiments, the MAEs in a age group is always lower when considering Gender or Ethnicity information, than without such considerations (using directly the selected features). When considering both Gender and Ethnicity in age estimation, the MAEs in each age group generally reach the lowest (marked in bold letter). Thus, by giving consideration to Gender and Ethnicity, we have successfully enhanced the age estimation performance for all the age groups.

Table 5.1 – Mean Absolute Errors (MAEs) of the fusion for different age groups under Expression-dependent and -independent settings.

Fusion/Age groups	≤ 20	(20,30]	(30,40]	> 40	All
Expression-dependent					
<i>Selected features</i>	3.99	2.11	6.48	24.50	3.45
<i>Gender-specific</i>	3.74	2.07	6.25	22.48	3.31
<i>Ethnicity-specific</i>	3.70	2.05	6.18	23.71	3.33
<i>Gender & Ethnicity-specific</i>	3.57	1.88	6.23	21.97	3.12
Expression-independent					
<i>Selected features</i>	3.90	2.34	6.69	23.83	3.82
<i>Gender-specific</i>	3.78	2.12	6.48	21.25	3.51
<i>Ethnicity-specific</i>	3.48	2.26	6.14	23.17	3.65
<i>Gender & Ethnicity-specific</i>	3.44	2.04	5.98	20.26	3.33
# of Subjects	185	246	20	15	466

5.3 HOW MUCH ARE GENDER, ETHNICITY AND AGE CORRELATED ?

The previous experimental section demonstrates that gender, ethnicity and age information are correlated, and their relationship are helpful in each others' recognition tasks. Following this, two questions rise up: (1) To how much extend are they correlated? and (2) How to benefit from their correlations in real-world like application where the ground truths of the other biometrics are unavailable? We address these questions in the following two sections.

Recall that with feature selection, we have obtained the salient subsets of features for Gender, Ethnicity and Age, for each of the four descriptions. Thus, we take these subsets as representations of the Gender, Ethnicity and Age information. With this, we explore two ways to quantify their mutual correlations. In the first way, we first represent a feature subset as a one dimensional vector with which the optimal class separation is obtained, then measure the correlation directly between such vectors. It provides a single value for each two biometrics, which represents their correlation in the *Decision Level*. To this end, we apply the Linear Discriminant Analysis (LDA) on each subset. The LDA method is a supervised dimensionality reduction method which projects the data into a subspace where optimal class separation could be obtained. The dimension of the projected subspace equals to the number of classes minus one. In our case, for each of the three biometrics, we have always two classes (Male and Female for Gender, Asian and Non-Asian for ethnicity, < 23 years and > 22 years groups for Age). Thus, after LDA projection, for each description, we get a one dimension vector for each attribute (Gender, Ethnicity and Age). In Fig. 5.4, we show the distribution of the projected LDA features for each biometric in each description, on the 466 earliest scans of FRGCv2. The x-axis shows the index of the scans, and the y axis shows the magnitude of the LDA projected features. Each row of the subfigures show the distributions of projected LDA features for a facial biometric, and each column of subfigures show the distributions in a facial description. In Fig. 5.4, clear separation of different classes is demonstrated for each biometric in each

description. It confirms that the feature selection method is able to keep the useful information. And also, the figure shows the LDA projected features are able to characterizing the three biometrics.

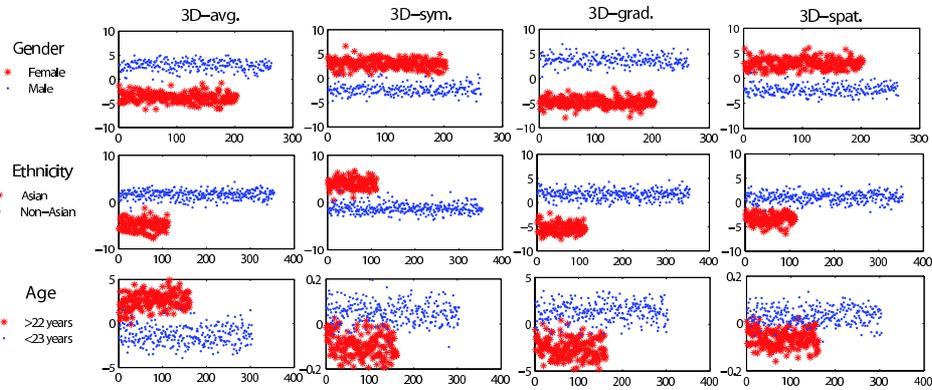


Figure 5.4 – Distribution of the projected LDA features for gender, ethnicity and age for each face description.

Table 5.2 – Pearson correlation coefficients between the Projected LDA features for the three biometrics.

	avg.	sym.	grad.	spat.
<i>Gender-Ethnicity</i>	0.1018	0.1025	0.1230	-0.1549
<i>Ethnicity-Age</i>	-0.6432	-0.6548	0.5719	0.6164
<i>Age-Gender</i>	-0.1006	-0.1050	0.1023	-0.1371

With the one dimension feature vector from LDA-projection, the correlation of these biometrics are able to be measured. We use the simple **Pearson correlation coefficient** between two vectors as the measurement. In Tab.5.2, we show the coefficients among these LDA projected features. The rows of the table show the correlation between each two attributes, and the columns show the correlation in different descriptions. Considering the absolute value, the correlation coefficients between Gender and Ethnicity are in the range of $[0.1, 0.16]$. It means that **Gender and Ethnicity are weakly correlated**. The absolute values of the correlation coefficients between Ethnicity and Age are in the range of $[0.57, 0.66]$. It shows that **Ethnicity and Age are strongly correlated**. The absolute values of the correlation coefficients between Age and Gender are in the range of $[0.1, 0.14]$. It shows that **Age and Gender are weakly correlated**.

The LDA projected features give a good summary of the information

for each biometric, and allow us to calculate the correlation of these biometrics in the **Decision level**. However, for a real face, these projected features convey no physic meaning. If we want to know their correlations in real face features, we need to examine in another way. Thus, we propose the second way to show the correlations among the three biometrics. The way is to perform biometric recognition with the selected features of the other biometrics. For example, we do Age estimation with selected features of Gender and Ethnicity. This will show their correlation in the **Feature Level** in terms of the recognition performance.

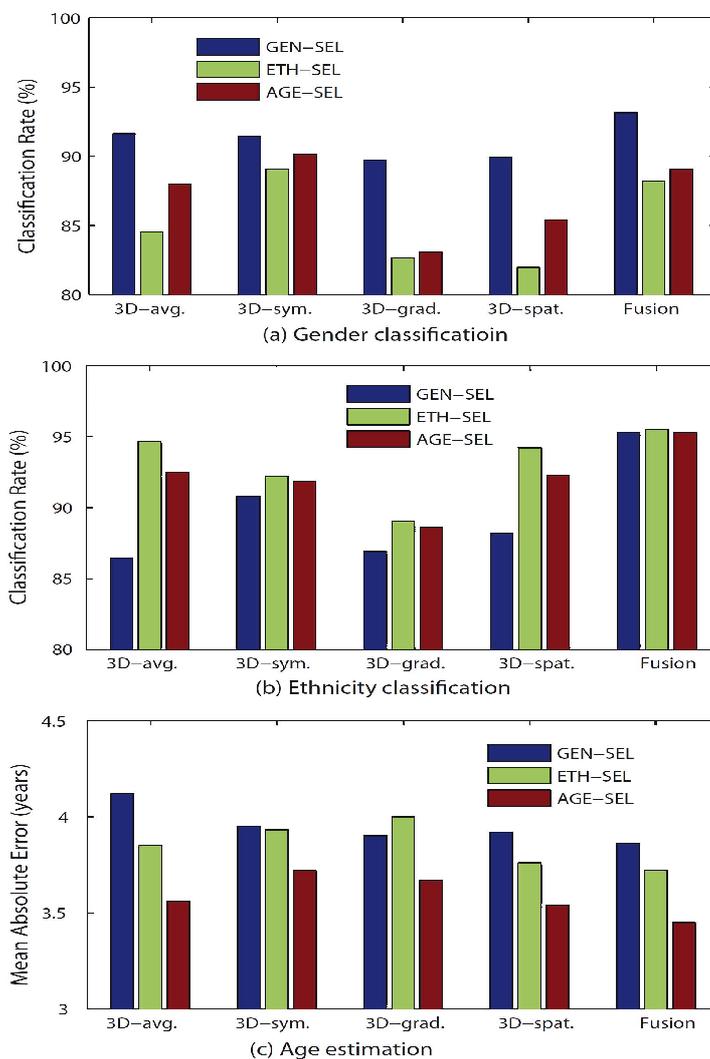


Figure 5.5 – Recognition results of the three biometrics using different selected features.

In Fig.5.5, we show the recognition results for the three biometrics with

different features, on the 466 earliest scans of FRGCv2 subjects. We find that for each facial biometric, the recognition performance is always the highest with its own selected features. For example, the selected features for gender work the best in Gender classification, as shown in Fig.5.5(a). While, the experiments with selected features for other biometrics also yield relatively good results. The Gender classification results are always $> 82\%$, the Ethnicity classification rates are always $> 86\%$, and the Age estimation MAEs are always < 4.2 years. These results show that in *Feature Level*, the Gender, Ethnicity and Age related features are strongly correlated in the face. Also, in panel (a) of Fig.5.5, the selected features for age perform better than selected features for ethnicity in Gender classification. **It indicates that the correlation between gender and age is stronger than the correlation between gender and ethnicity.** In panel (b) of Fig.5.5, the selected features for age perform better than selected features for gender in Ethnicity classification. **It means that the correlation between ethnicity and age is stronger than the correlation between ethnicity and gender.** In the panel (c) of Fig.5.5, the selected features for ethnicity generally perform better than selected features for gender in Age estimation. **It means that the correlation between ethnicity and age is stronger than the correlation between gender and age.** Thus, in Feature Level, the correlation between Ethnicity and Age is stronger than the correlation between Gender and Age, and further stronger than the correlation between Gender and Ethnicity. **Thus, both at the *Feature Level* and the *Decision Level*, we find out that the correlation between Ethnicity and Age is the strongest.**

5.4 HOW TO BENEFIT FROM THEIR CORRELATION IN REAL-WORLD LIKE APPLICATIONS ?

In real-world like applications, we do not have the ground truth of Gender, Ethnicity and Age of the testing instances. Benefited from the effective recognition performance in the previous section, we can use the recognition results as the gender, ethnicity and age information in the experiments. Thus, instead of using the ground truth, we enroll gender,

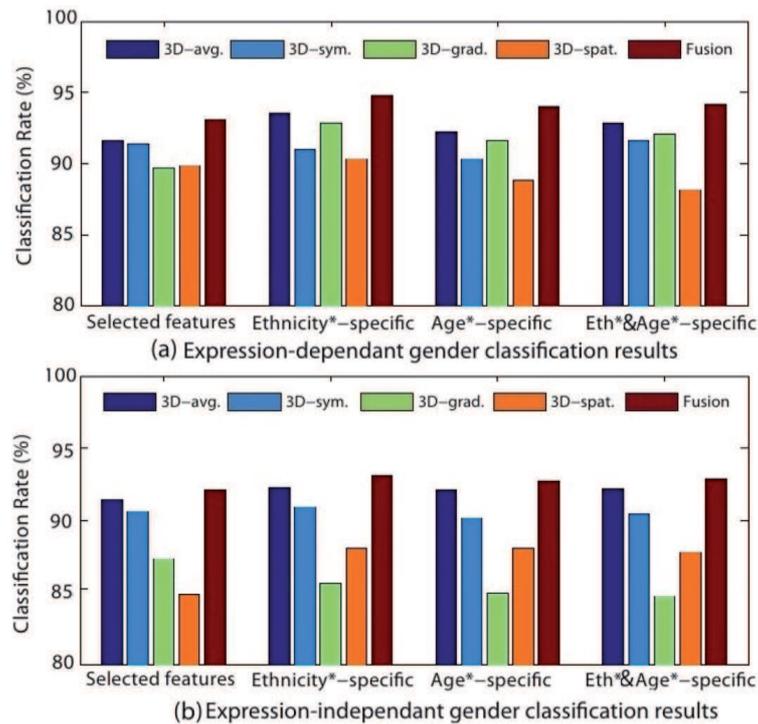


Figure 5.6 – Gender classification using automatic recognition results of ethnicity and age.

ethnicity and age information with the predicted gender, ethnicity and age labels given in the previous recognition tasks in the *feature Selection* setting with the *Fusion* description. The predicted information are tagged with * as *Gender**, *Ethnicity** and *Age**. The recognition results are reported in Figure 5.6-5.8.

In Fig.5.6 (a) and (b), the gender classification results are presented. In the Expression-dependent experiments, except for the 3D-spat. description, the gender classification results are always higher when considering ethnicity and age information, than without such consideration (in the *Selected features* setting). In the Expression-independent experiments, except for the 3D-grad. description, the gender classification results are always higher when considering ethnicity and age information. With the Fusion description in the *Ethnicity*-specific* setting, we achieve 94.85% Gender classification rate in the Expression-dependent experiments, and 93.08% Gender classification rate in the Expression-independent experiments.

The experimental results of ethnicity classification are presented in Fig.5.7 (c) and (d). In both the Expression-dependent and the Expression-

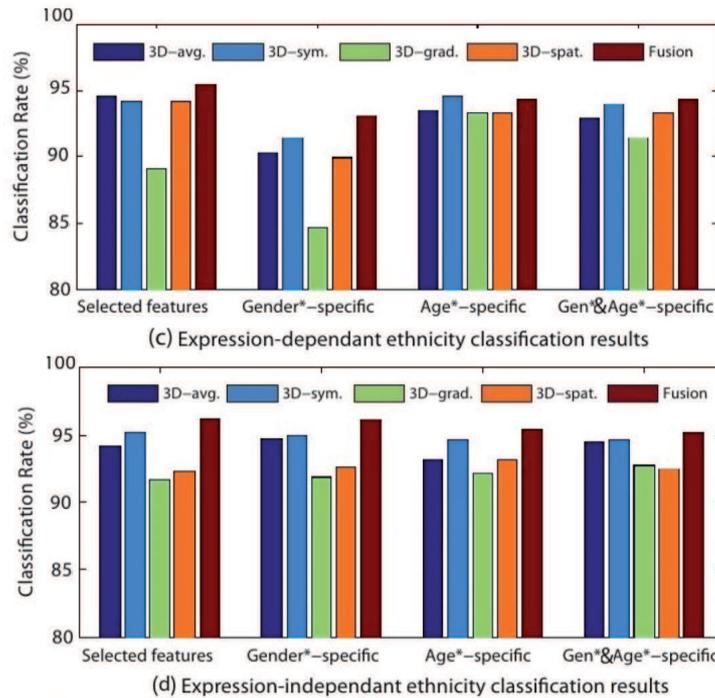


Figure 5.7 – Ethnicity classification using automatic recognition results of gender and age.

independent experiments, the results considering Gender and Age information are comparable or slightly lower, than without such consideration (in the *Selected features* setting). With the Fusion description, we achieve 94.42% Ethnicity classification rate in the *Age*-specific* and the *Gen*&Age*-specific* settings in the Expression-dependent experiments, and 96.18% Ethnicity classification rate in the *Gen*-specific* settings in the Expression-independent experiments.

For Age estimation, the results are shown in Fig.5.8 (e) and (f). Compared to the *Selected features* setting, the MAEs are significantly reduced when considering Gender and Ethnicity information in both the Expression-dependent and the the Expression-independent experiments. With the Fusion description in the *Gen*&Eth*-specific* setting, we achieve a MAE of 3.13 years in the Expression-dependent experiments, and 3.62 years MAE in the Expression-independent experiments. The result also confirm the previous finding that the combination of gender and ethnicity information gives the strongest enhancement of age estimation performance. Thus, in summary of these experiments, it is clear that, for

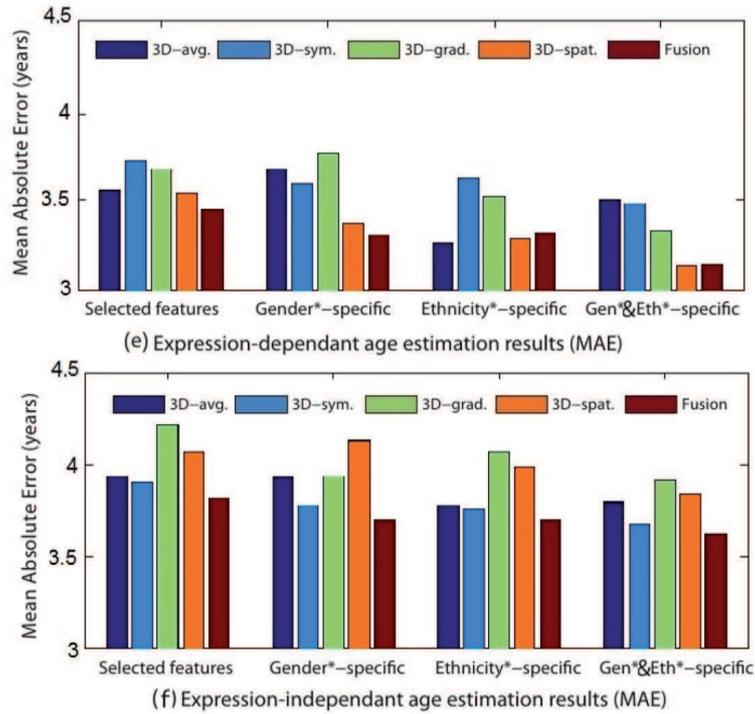


Figure 5.8 – *Age estimation using automatic recognition results of gender and ethnicity.*

Gender and Age recognition, we have obtained higher performance with the automatically recognized information, than without using these information. For Ethnicity recognition, since the training scans are reduced significantly and the classification rate reaches as high as 94.42% (Expression-dependent) and 96.18% (Expression-independent) when considering gender or age information, we still think that using the automatically recognized information of gender and age is a strategic solution in real world-like application.

5.5 CONCLUSION

In this chapter, we have made the first thorough examination in the literature for the correlation among gender, ethnicity and age within the 3D face. Experimental results confirm that gender, ethnicity and age are correlated with each other in 3D face, and their correlations can be useful in each others' recognition tasks. The correlation between ethnicity and age is recognized as the strongest among their correlations. With the

FRGCv2 dataset, we have demonstrated significant competence in related recognition tasks. In comparison with the performances using directly the selected features, we have achieved better results with the *biometric-specific* settings in both the Expression-Dependent and Expression-Independent settings. With the 466 earliest scans of FRGCv2, we achieve 94.64% recognition rate for gender classification, 95.71% recognition rate for ethnicity classification, and 3.08 years MAE for age estimation. With all the 4007 scans in FRGCv2, we gain 93.11% recognition rate for gender classification, 96.60% recognition rate for ethnicity classification, and 3.33 years MAE for age estimation. We have also demonstrated that using the automatic recognized information of the other soft-biometrics, the recognition performance can still be enhanced. By giving consideration of their correlations, the demand of the computational resources is significantly reduced. This indicates a possible solution for real-time application in low computational resource scenarios, such as in the mobiles.

CONCLUSION

6

In this thesis, we aimed at learning geometric features from 3D facial surfaces for Gender, Ethnicity and Age recognition. To this end, we proposed to use a set of facial morphology cues which are closely related to these facial soft-biometrics, namely the face Averageness, the bilateral Symmetry, the global Spatial configuration and the local shape gradient information. Through the Riemannian shape analysis of elastic open curves, we extracted four types of Dense Scalar Field (DSF) features on each point of the face, in consideration of each type of the morphology cues. Then, we explore the extracted features for facial soft-biometric recognition on the FRGCv2 dataset, in two experimental contexts. First, we explored the usage of these features with Random Forest for recognizing Gender, Ethnicity and Age individually. In this context, we experimented directly with the 466 earliest scans of FRGCv2 dataset under the Expression-Dependent setting, and experimented with the whole 4007 scans of FRGCv2 under the Expression-Independent setting. To deal with the high dimensionality of the DSF features, we proposed to use the Principal Component Analysis (PCA) and the Correlation-based Feature Selection (CFS) for feature dimensionality reduction. We also proposed a fusion method which concatenates the four descriptions. Experimental results from both the Expression-Dependent and Expression-Independent settings show that the proposed DSF descriptions are effective in Gender, Ethnicity and Age estimation. It justifies the close relationship between the facial morphology cues and the facial soft-biometrics. Results also revealed that the CFS method outperforms the PCA in the related tasks. After Correlation-based Feature Selection, the feature dimensionality reduces to 200-400, and the performance of Gender, Ethnicity and Age recognition improved significantly. The fusion method always demonstrated better results than each individual description in all the related experiments. It means that the proposed DSF features are complimentary in recognition of these facial soft-biometrics. Furtherly, it means the underlying facial morphology cues are complimentary in revealing the Gender, Ethnicity and Age in 3D faces.

In the second context, we recognized these soft-biometrics jointly in consideration of their correlations. With *biometric-specific* experimental settings which work on specific demographic group of the population, we always obtained better recognition performance for these soft-biometrics, in comparison to the results when recognizing each of them individually. It means that the Gender patterns, the Ethnicity patterns and the Age patterns are correlated with each other, and their correlations are useful in each others' recognition tasks. With both the *decision-level* and the *feature-level* level analysis, we discovered that the correlation between age and ethnicity is the strongest among their correlations. We also demonstrated that in real-world like applications where the ground truth of the query instance is not available, the automatic recognized information can be used to perform the *biometric-specific* experiments. Following this idea, the correlations demonstrated the usage in each others' recognition task, in the real-world like application.

The proposed approach has also some limitations. First, it relies on the near-frontal faces to detect the nose tip for feature extraction. It works well on the FRGCv2 dataset, because the scans in FRGCv2 dataset are all near-frontal. For posed faces, this proposed approach lacks the capability to detect accurately the nose tip. Secondly, the experiments are based on high resolution 3D scans in the FRGCv2. In real world application, the 3D scans are usually captured by low resolution scanners, such as the Kinect scanner. Thus, the effectiveness of the proposed approach on the low resolution 3D scans needs to be tested. Thirdly, the presented approach only concerns the 3D shape of the face. The 2D texture information of the face has been filtered out arbitrarily. Thus, one possible enhancement of the proposed approach is to combine 2D and 3D information together in the related recognition tasks.

BIBLIOGRAPHY

- [1] The main differences between male and female faces? In *www.virtualffs.co.uk*. (Cited pages 4, 8, 9, 13, 26, 35, 38, 40, and 52.)
- [2] Male skin vs. female skin. In *www.faroche.com*. (Cited pages 22 and 82.)
- [3] Meyers konversationslexikon 1985-1990. (Cited page 15.)
- [4] Population variation in the front view of the face. In *www.femininebeauty.info*. (Cited pages 4, 8, 27, 35, 38, and 52.)
- [5] O'toole Alice J, Vetter Thomas, Volz Harald, and Salter Elizabeth M. Three-dimensional caricatures of human heads: Distinctiveness and the perception of facial age. In *PERCEPTION*, volume 26, pages 719–732. Citeseer, 1997. (Cited pages 27 and 35.)
- [6] Jennifer Alphonse, Jennifer Cox, Jill Clarke, Philip Schluter, and Andrew McLennan. The effect of ethnicity on 2d and 3d frontomaxillary facial angle measurement in the first trimester. In *Obstetrics and Gynecology International*, 2013. (Cited pages 4, 8, 27, 35, and 52.)
- [7] K. Jain Anil, C. Dass Sarat, Nandakumar Karthik, and N Karthik. Soft biometric traits for personal recognition systems. In *Proceedings of International Conference on Biometric Authentication, Hong Kong*, pages 731–738. Springer, 2004. (Cited page 2.)
- [8] Samal Ashok, Subramani Vanitha, and Marx David. Analysis of sexual dimorphism in human face. *Journal of Visual Communication and Image Representation*, 18:453–463, 2007. (Cited page 27.)
- [9] Lahoucine Ballihi, Boulbaba Ben Amor, Mohamed Daoudi, Anuj Srivastava, and Driss Aboutajdine. Boosting 3D-geometric features for

- efficient face recognition and gender classification. In *Information Forensics and Security, IEEE Transactions on*, volume 7, pages 1766–1779. IEEE, 2012. (Cited pages 13, 47, 77, and 78.)
- [10] Jean-Yves Baudouin and Mathieu Gallay. Is face distinctiveness gender based? *Journal of Experimental Psychology: Human Perception and Performance*, 32:789, 2006. (Cited pages 26 and 35.)
- [11] Juan Bekios-Calfa, Jose M Buenaposada, and Luis Baumela. Revisiting linear discriminant techniques in gender recognition. In *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, volume 33, pages 858–864. IEEE, 2011. (Cited page 11.)
- [12] Boulbaba Ben Amor, Hassen Drira, Stefano Berretti, Mohamed Daoudi, and Anuj Srivastava. 4d facial expression recognition by learning geometric deformations. In *IEEE Transactions on Cybernetics*, accepted for publication, February 2014. (Cited pages 28 and 29.)
- [13] Leo Breiman. Random forests. In *Machine learning*, volume 45, pages 5–32. Springer, 2001. (Cited page 47.)
- [14] Wang C, Huang D, Wang Y, and Zhang G. Facial image-based gender classification using local circular patterns. In *International Conference on Pattern Recognition*, pages 2432–2435. IEEE, 2012. (Cited page 11.)
- [15] Coon Carleton S. The origin of races. 1962. (Cited page 15.)
- [16] Morrison Clinton S, Phillips Benjamin Z, Chang Johnny T, Sullivan Stephen R, and Taylor Helena OB. The relationship between age and facial asymmetry. In <http://meeting.nesps.org/2011/80.cgi>, 2011. (Cited pages 17, 27, and 34.)
- [17] Timothy F Cootes, Gareth J Edwards, and Christopher J Taylor. Active appearance models. In *Computer Vision-ECCV'98*, pages 484–498. Springer, 1998. (Cited page 20.)
- [18] Antitza Dantcheva, Jean-Luc Dugelay, and Petros Elia. Person recognition using a bag of facial soft biometrics (bofsb). In *Multimedia*

- Signal Processing (MMSP), 2010 IEEE International Workshop on*, pages 511–516. IEEE, 2010. (Cited page 2.)
- [19] Antitza Dantcheva, Carmelo Velardo, Angela D’angelo, and Jean-Luc Dugelay. Bag of soft biometrics for person identification. *Multimedia Tools and Applications*, 51:739–777, 2011. (Cited pages 2 and 3.)
- [20] Kenneth A Deffenbacher, Thomas Vetter, John Johanson, and Alice J O’Toole. Facial aging, attractiveness, and distinctiveness. In *PERCEPTION*, volume 27, pages 1233–1244. PION LTD, 1998. (Cited pages 27 and 35.)
- [21] Meltem Demirkus, Kshitiz Garg, and Sadiye Guler. Automated person categorization for video surveillance using soft biometrics. In *SPIE Defense, Security, and Sensing*, pages 76670P–76670P. International Society for Optics and Photonics, 2010. (Cited pages 15 and 16.)
- [22] Huaxiong Ding, Di Huang, Yunhong Wang, and Liming Chen. Facial ethnicity classification based on boosted local texture and shape descriptions. In *10th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition, FG 2013*, pages 1–6, 2013. (Cited page 17.)
- [23] Hassen Drira, B Ben Amor, Mohamed Daoudi, Anuj Srivastava, and Stefano Berretti. 3d dynamic expression recognition based on a novel deformation vector field and random forest. In *Pattern Recognition (ICPR), 2012 21st International Conference on*, pages 1104–1107. IEEE, 2012. (Cited pages 28 and 29.)
- [24] Hassen Drira, Boulbaba Ben Amor, Anuj Srivastava, Mohamed Daoudi, and Rim Slama. 3d face recognition under expressions, occlusions, and pose variations. In *IEEE Transactions on Pattern Analysis and Machine Intelligence*, volume 35, pages 2270–2283, 2013. (Cited pages 28, 29, and 47.)
- [25] Lili Du, Ziqing Zhuang, Hongyu Guan, Jingcai Xing, Xianzhi Tang, Limin Wang, Zhenglun Wang, Haijiao Wang, Yuewei Liu, Wenjin

- Su, et al. Head-and-face anthropometric survey of chinese workers. In *Annals of occupational hygiene*, volume 52, pages 773–782. BOHS, 2008. (Cited pages 27 and 35.)
- [26] Donald H Enlow. A morphogenetic analysis of facial growth. In *American journal of orthodontics*, volume 52, pages 283–299. Elsevier, 1966. (Cited pages 4, 8, 27, and 35.)
- [27] Ilker Ercan, Senem Turan Ozdemir, Abdullah Etoz, Deniz Sigirli, R Shane Tubbs, Marios Loukas, and Ibrahim Guney. Facial asymmetry in young healthy subjects evaluated by statistical shape analysis. *Journal of anatomy*, 213:663–669, 2008. (Cited page 39.)
- [28] Giovanna Farinella and Jean-Luc Dugelay. Demographic classification: Do gender and ethnicity affect each other? In *International Conference on Informatics Electronics and Vision (ICIEV)*, pages 383–390, 2012. (Cited pages 11, 16, 22, and 82.)
- [29] Leslie G Farkas and Gwynne Cheung. Facial asymmetry in healthy north american caucasians: An anthropometrical study. *The Angle orthodontist*, 51:70–77, 1981. (Cited page 39.)
- [30] Leslie G Farkas and Stephen A Schendel. Anthropometry of the head and face. In *American Journal of Orthodontics and Dentofacial Orthopedics*, volume 107, pages 112–112, 1995. (Cited pages 4 and 8.)
- [31] Virgilio F Ferrario, Chiarella Sforza, Veronica Ciusa, Claudia Dellavia, and Gianluca M Tartaglia. The effect of sex and age on facial asymmetry in healthy subjects: a cross-sectional study from adolescence to mid-adulthood. *Journal of oral and maxillofacial surgery*, 59:382–388, 2001. (Cited page 39.)
- [32] Yun Fu, Guodong Guo, and Thomas S Huang. Age synthesis and estimation via faces: A survey. In *IEEE Transactions on Pattern Analysis and Machine Intelligence*, volume 32, pages 1955–1976, 2010. (Cited pages 2, 3, 18, and 21.)
- [33] Wei Gao and Haizhou Ai. Face gender classification on consumer

- images in a multiethnic environment. In *Advances in Biometrics*, pages 169–178, 2009. (Cited pages 11, 22, and 82.)
- [34] Xin Geng, Zhi-Hua Zhou, and Kate Smith-Miles. Automatic age estimation based on facial aging patterns. In *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, volume 29, pages 2234–2240. IEEE, 2007. (Cited pages 19, 21, and 53.)
- [35] Xin Geng, Zhi-Hua Zhou, Yu Zhang, Gang Li, and Honghua Dai. Learning from facial aging patterns for automatic age estimation. In *Proceedings of the 14th annual ACM international conference on Multimedia*, pages 307–316. ACM, 2006. (Cited page 21.)
- [36] Syed Zulqarnain Gilani, Faisal Shafait, and Ajmal Mian. Biologically significant facial landmarks: How significant are they for gender classification? In *Digital Image Computing: Techniques and Applications (DICTA), 2013 International Conference on*, pages 1–8. IEEE, 2013. (Cited pages 14, 47, 77, and 78.)
- [37] Vignali Guillaume, Hill Harold, and Vatikiotis Bateson Eric. Linking the structure and perception of 3d faces: Gender, ethnicity ,and expressive posture. In *International Conference on Audio-Visual Speech Processing (AVSP)*, 2003. (Cited pages 22 and 82.)
- [38] Guodong Guo, Yun Fu, Charles R Dyer, and Thomas S Huang. Image-based human age estimation by manifold learning and locally adjusted robust regression. In *Image Processing, IEEE Transactions on*, volume 17, pages 1178–1188. IEEE, 2008. (Cited pages 20, 21, 22, 23, and 82.)
- [39] Guodong Guo, Yun Fu, Thomas S Huang, and Charles R Dyer. Locally adjusted robust regression for human age estimation. In *Applications of Computer Vision, 2008. WACV 2008. IEEE Workshop on*, pages 1–6. IEEE, 2008. (Cited pages 18, 20, and 21.)
- [40] Guodong Guo and Guowang Mu. Simultaneous dimensionality reduction and human age estimation via kernel partial least squares

- regression. In *Computer Vision and Pattern Recognition*, pages 657–664. IEEE, 2011. (Cited page 23.)
- [41] Guodong Guo and Guowang Mu. Joint estimation of age, gender and ethnicity: Cca vs. pls. In *Automatic Face and Gesture Recognition (FG)*, pages 1–6. IEEE, 2013. (Cited page 23.)
- [42] Guodong Guo, Guowang Mu, Yun Fu, and Thomas S Huang. Human age estimation using bio-inspired features. In *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*, pages 112–119. IEEE, 2009. (Cited pages 18 and 20.)
- [43] Guo Guodong and Mu Guowang. Human age estimation: What is the influence across race and gender? In *Computer Vision and Pattern Recognition Workshops (CVPRW)*, pages 71–78, June 2010. (Cited pages 23 and 82.)
- [44] Srinivas Gutta, Harry Wechsler, and P Jonathon Phillips. Gender and ethnic classification of face images. In *IEEE International Conference on Automatic Face and Gesture Recognition*, pages 194–199. IEEE, 1998. (Cited pages 15 and 16.)
- [45] Mark A. Hall. Correlation-based feature subset selection for machine learning. In *PhD thesis, Department of Computer Science, University of Waikato*, 1999. (Cited page 49.)
- [46] Hu Han, Charles Otto, and Anil K Jain. Age estimation from face images: Human vs. machine performance. In *The 6th IAPR International Conference on Biometrics (ICB)*, 2013. (Cited pages 18 and 21.)
- [47] Xia Han, Hassan Ugail, and Ian Palmer. Gender classification based on 3D face geometry features using svm. In *International Conference on CyberWorlds*, pages 114–118, 2009. (Cited page 13.)
- [48] Jun-ichiro Hayashi, Mamoru Yasumoto, Hideaki Ito, and Hiroyasu Koshimizu. Age and gender estimation based on wrinkle texture and color of facial images. In *Pattern Recognition, 2002. Proceedings. 16th International Conference on*, volume 1, pages 405–408. IEEE, 2002. (Cited page 28.)

- [49] Harold Hill, Vicki Bruce, and Shigeru Akamatsu. Perceiving the sex and race of faces: The role of shape and colour. In *Proceedings Of The Royal Society Of London Series B Biological Sciences*, volume 261(1362), pages 367–373, 1995. (Cited page 5.)
- [50] Yuan Hu, Jingqi Yan, and Pengfei Shi. A fusion-based method for 3D facial gender classification. In *Computer and Automation Engineering (ICCAE)*, volume 5, pages 369–372. IEEE, 2010. (Cited page 13.)
- [51] Yuxiao Hu, Yun Fu, Usman Tariq, and Thomas S Huang. Subjective experiments on gender and ethnicity recognition from different face representations. In *Advances in Multimedia Modeling*, volume 5916, pages 66–75, 2010. (Cited page 5.)
- [52] Di Huang, Huaxiong Ding, Chen Wang, Yunhong Wang, Guang-peng Zhang, and Liming Chen. Local circular patterns for multi-modal facial gender and ethnicity classification. *Image and Vision Computing*, 2014. (Cited pages 15, 17, 77, and 78.)
- [53] W Stuart Hunter and Stanley M Garn. Disproportionate sexual dimorphism in the human face. In *American Journal of Physical Anthropology*, volume 36, pages 133–138. Wiley online Library, 1972. (Cited pages 9, 26, 35, 40, and 52.)
- [54] Tri Huynh, Rui Min, and Jean-Luc Dugelay. An efficient lbp-based descriptor for facial depth images applied to gender recognition using rgb-d face data. In *ACCV 2012, Workshop on Computer Vision with Local Binary Pattern Variants*, pages 133–145. Springer, 2013. (Cited page 15.)
- [55] Anil K Jain, Sarat C Dass, and Karthik Nandakumar. Can soft biometric traits assist user recognition? In *Defense and Security*, pages 561–572. International Society for Optics and Photonics, 2004. (Cited page 2.)
- [56] Anil K Jain, Sarat C Dass, and Karthik Nandakumar. Soft biometric traits for personal recognition systems. In *Biometric Authentication*, pages 731–738. Springer, 2004. (Cited page 2.)

- [57] Anil K Jain and Unsang Park. Facial marks: Soft biometric for face recognition. In *Image Processing (ICIP), 2009 16th IEEE International Conference on*, pages 37–40. IEEE, 2009. (Cited page 3.)
- [58] Tetsuro Shuler John. Facial sexual dimorphism and judgments of personality: A literature review. In *Issues*, volume 6, 2012. (Cited page 9.)
- [59] Benedict C Jones, Lisa M DeBruine, and Anthony C Little. The role of symmetry in attraction to average faces. *Perception & psychophysics*, 69:1273–1277, 2007. (Cited page 26.)
- [60] Young Joseph W. Head and face anthropometry of adult us civilians. 1993. (Cited pages 27 and 35.)
- [61] Ueki Kazuya, Sugiyama Masashi, and Ihara Yasuyuki. Perceived age estimation under lighting condition change by covariate shift adaptation. In *Pattern Recognition (ICPR), 2010 20th International Conference on*, pages 3400–3403. IEEE, 2010. (Cited pages 22, 23, and 82.)
- [62] Masashi Komori, Satoru Kawamura, and Shigekazu Ishihara. Effect of averageness and sexual dimorphism on the judgment of facial attractiveness. *Vision research*, 49:862–869, 2009. (Cited page 26.)
- [63] Neeraj Kumar, Alexander C Berg, Peter N Belhumeur, and Shree K Nayar. Describable visual attributes for face verification and image search. In *IEEE Transactions on Pattern Analysis and Machine Intelligence*, volume 33, pages 1962–1977, 2008. (Cited page 10.)
- [64] NS Lakshmiprabha, J Bhattacharya, and S Majumder. Age estimation using gender information. In *Computer Networks and Intelligent Computing*, pages 211–216. Springer, 2011. (Cited pages 22, 23, and 82.)
- [65] Andreas Lanitis. On the significance of different facial parts for automatic age estimation. In *Digital Signal Processing, 14th International Conference on*, volume 2, pages 1027–1030. IEEE, 2002. (Cited pages 27, 28, and 35.)

- [66] Andreas Lanitis. Facial age estimation. In *Scholarpedia*, volume 5, page 9701, 2010. (Cited page 19.)
- [67] Andreas Lanitis, Chrisina Draganova, and Chris Christodoulou. Comparing different classifiers for automatic age estimation. In *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, volume 34, pages 621–628. IEEE, 2004. (Cited pages 20 and 21.)
- [68] Andreas Lanitis, Christopher J. Taylor, and Timothy F. Cootes. Toward automatic simulation of aging effects on face images. In *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, volume 24, pages 442–455. IEEE, 2002. (Cited pages 18, 20, 21, and 22.)
- [69] Farkas Leslie G, Katic Marko J, and Forrest Christopher R. International anthropometric study of facial morphology in various ethnic groups/races. In *Journal of Craniofacial Surgery*, volume 16, pages 615–646, 2005. (Cited pages 4, 8, 12, 27, 35, 38, and 52.)
- [70] Changsheng Li, Qingshan Liu, Jing Liu, and Hanqing Lu. Learning ordinal discriminative features for age estimation. In *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on*, pages 2570–2577. IEEE, 2012. (Cited pages 20 and 22.)
- [71] Anthony C Little, Benedict C Jones, Lisa M DeBruine, and David R Feinberg. Symmetry and sexual dimorphism in human faces: inter-related preferences suggest both signal quality. *Behavioral Ecology*, 19:902–908, 2008. (Cited pages 26 and 28.)
- [72] Anthony C Little, Benedict C Jones, Corri Waitt, Bernard P Tiddeman, David R Feinberg, David I Perrett, Coren L Apicella, and Frank W Marlowe. Symmetry is related to sexual dimorphism in faces: Data across culture and species. In *PLoS ONE*, volume 3, page e2106. Public Library of Science, 05 2008. (Cited pages 26, 27, and 34.)
- [73] Yanxi Liu and Jeff Palmer. A quantified study of facial asymmetry in 3d faces. In *Analysis and Modeling of Faces and Gestures*, pages 222–229, 2003. (Cited pages 12 and 52.)

- [74] Xiaoguang Lu, Hong Chen, and Anil K Jain. Multimodal facial gender and ethnicity identification. In *Advances in Biometrics*, pages 554–561. Springer, 2005. (Cited pages 14 and 17.)
- [75] Xiaoguang Lu and Anil K Jain. Ethnicity identification from face images. In *International Symposium on Defense and Security : Biometric Technology for Human Identification*, pages 114–123, 2004. (Cited pages 4, 8, 15, and 16.)
- [76] Erno Mäkinen and Roope Raisamo. An experimental comparison of gender classification methods. In *Pattern Recognition Letters*, volume 29, pages 1544–1556. Elsevier, 2008. (Cited page 9.)
- [77] Choon Boon Ng, Yong Haur Tay, and Bok Min Goi. Vision-based human gender recognition: a survey. In *CoRR*, 2012. (Cited page 3.)
- [78] Unsang Park, Yiyong Tong, and Anil K Jain. Age-invariant face recognition. In *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, volume 32, pages 947–954. IEEE, 2010. (Cited page 18.)
- [79] Claudio Perez, Juan Tapia, Pablo Estevez, and Claudio Held. Gender classification from face images using mutual information and feature fusion. In *International Journal of Optomechatronics*, volume 6, pages 92–119, 2012. (Cited page 12.)
- [80] P Jonathon Phillips, Patrick J Flynn, Todd Scruggs, Kevin W Bowyer, Jin Chang, Kevin Hoffman, Joe Marques, Jaesik Min, and William Worek. Overview of the face recognition grand challenge. In *Computer Vision and Pattern Recognition*, volume 1, pages 947–954. IEEE, 2005. (Cited pages 29 and 46.)
- [81] Narayanan Ramanathan, Rama Chellappa, and Soma Biswas. Computational methods for modeling facial aging: A survey. In *Journal of Visual Languages & Computing*, volume 20, pages 131–144. Elsevier, 2009. (Cited pages 4, 8, 19, 22, 23, and 82.)
- [82] Gillian Rhodes, Leslie A Zebrowitz, Alison Clark, S Michael Kalick, Amy Hightower, and Ryan McKay. Do facial averageness and sym-

- metry signal health? *Evolution and Human Behavior*, 22:31–46, 2001. (Cited page 26.)
- [83] Henry Rhodes. *Alphonse Bertillon: Father of Scientific Detection*. Abelard-Schuman, 1956. (Cited page 2.)
- [84] Matthew G Rhodes. Age estimation of faces: a review. In *Applied Cognitive Psychology*, volume 23, pages 1–12. Wiley Online Library, 2009. (Cited pages 4, 8, 17, 18, and 19.)
- [85] Kohavi Ron. Wrappers for performance enhancement and oblivious decision graphs. In *PhD thesis, Stanford University*, 1995. (Cited page 49.)
- [86] Caifeng Shan. Learning local binary patterns for gender classification on real-world face images. In *Pattern Recognition Letters*, volume 33, pages 431–437, 2012. (Cited page 11.)
- [87] Finlay G Smith, Benedict C Jones, Lisa M DeBruine, and Anthony C Little. Interactions between masculinity–femininity and apparent health in face preferences. *Behavioral Ecology*, 20:441–445, 2009. (Cited page 26.)
- [88] Anuj Srivastava, Eric Klassen, Shantanu H Joshi, and Ian H Jermyn. Shape analysis of elastic curves in euclidean spaces. In *IEEE Transactions on Pattern Analysis and Machine Intelligence*, volume 33, pages 1415–1428, 2011. (Cited pages 29, 30, and 31.)
- [89] WG Steven and T Randy. Facial masculinity and fluctuating asymmetry. *Evolution and Human Behavior*, 24:231–241, 2003. (Cited pages 27 and 34.)
- [90] Jinli Suo, Song-Chun Zhu, Shiguang Shan, and Xilin Chen. A compositional and dynamic model for face aging. In *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, volume 32, pages 385–401. IEEE, 2010. (Cited pages 21, 27, and 35.)
- [91] George Toderici, Sean M O’Malley, George Passalis, Theoharis Theoharis, and Ioannis A Kakadiaris. Ethnicity- and gender-based sub-

- ject retrieval using 3-D face-recognition techniques. In *International Journal of Computer Vision*, volume 89, pages 382–391. Springer, 2010. (Cited pages 13, 16, 77, and 78.)
- [92] Bruce Vicki, Burton A Mike, Hanna Elias, Healey Pat, Mason Oli, Coombes Anne, Fright Rick, and Linney Alf. Sex discrimination: How do we tell the difference between male and female faces? In *Perception*, volume 22(2), pages 131–152, 1993. (Cited pages 4, 8, 9, 26, 35, 38, 40, and 52.)
- [93] Xiaolong Wang and Chandra Kambhamettu. Gender classification of depth images based on shape and texture analysis. In *Global Conference on Signal and Information Processing (GlobalSIP)*, pages 1077–1080. IEEE, 2013. (Cited pages 77 and 78.)
- [94] Svante Wold, Kim Esbensen, and Paul Geladi. Principal component analysis. In *Chemometrics and intelligent laboratory systems*, volume 2, pages 37–52. Elsevier, 1987. (Cited page 49.)
- [95] Jing Wu, William AP Smith, and Edwin R Hancock. Gender classification using shape from shading. In *BMVC*, pages 1–10, 2007. (Cited page 14.)
- [96] Tao Wu, Pavan Turaga, and Rama Chellappa. Age estimation and face verification across aging using landmarks. In *Information Forensics and Security, IEEE Transactions on*, volume 7, pages 1780–1788. IEEE, 2012. (Cited page 20.)
- [97] Baiqiang Xia, Boulbaba Ben Amor, Mohamed Daoudi, and Hassen Drira. Can 3d shape of the face reveal your age? In *International Conference on Computer Vision Theory and Applications*, pages 5–13, 2014. (Cited page 77.)
- [98] Baiqiang Xia, Boulbaba Ben Amor, Hassen Drira, Mohamed Daoudi, and Lahoucine Ballihi. Gender and 3D facial symmetry: What’s the relationship? In *IEEE Conference on Automatic Face and Gesture Recognition*, 2013. (Cited page 77.)

- [99] Wankou Yang, Cuixian Chen, Karl Ricanek, and Changyin Sun. Gender classification via global-local features fusion. In *Biometric Recognition*, volume 7098, pages 214–220, 2011. (Cited page 10.)
- [100] Zhiguang Yang and Haizhou Ai. Demographic classification with local binary patterns. In *Advances in Biometrics*, pages 464–473. Springer, 2007. (Cited pages 15 and 16.)
- [101] Juha Ylioinas, Abdenour Hadid, and Matti Pietikäinen. Combining contrast information and local binary patterns for gender classification. In *Image Analysis*, volume 6688, pages 676–686, 2011. (Cited page 10.)
- [102] and Yoshiro Suzuki Yukio, Shirakabe and Lam Samuel M. A new paradigm for the aging asian face. In *Aesthetic Plastic Surgery*, volume 27, pages 397–402, 2003. (Cited pages 22 and 82.)
- [103] LA Zebrowitz. Reading faces: Window to the soul? Westview Press, 1997. (Cited page 2.)
- [104] Cheng Zhong, Zhenan Sun, and Tieniu Tan. Fuzzy 3d face ethnicity categorization. In *Advances in Biometrics*, pages 386–393. 2009. (Cited pages 16, 47, and 78.)
- [105] Ziqing Zhuang and Bruce Bradtmiller. Head and face anthropometric survey of us respirator users. In *Journal of occupational and environmental hygiene*, volume 2, pages 567–576. Taylor & Francis, 2005. (Cited pages 27 and 35.)
- [106] Ziqing Zhuang, Douglas Landsittel, Stacey Benson, Raymond Roberge, and Ronald Shaffer. Facial anthropometric differences among gender, ethnicity, and age groups. *Annals of Occupational Hygiene*, page meq007, 2010. (Cited pages 4, 8, 9, 22, 26, 27, 35, 40, 52, and 82.)

Résumé La reconnaissance des biométries douces (genre, âge, etc.) trouve ses applications dans plusieurs domaines. Les approches proposées se basent sur l'analyse de l'apparence (images 2D), très sensibles aux changements de la pose et à l'illumination, et surtout pauvre en descriptions morphologiques. Dans cette thèse, nous proposons d'exploiter la forme 3D du visage. Basée sur une approche Riemannienne d'analyse de formes 3D, nous introduisons quatre descriptions denses à savoir: la symétrie bilatérale, la moyenneté, la configuration spatiale et les variations locales de sa forme. Les évaluations faites sur la base FRGCv2 montrent que l'approche proposée est capable de reconnaître des biométries douces. A notre connaissance, c'est la première étude menée sur l'estimation de l'âge, et c'est aussi la première étude qui propose d'explorer les corrélations entre les attributs faciaux, à partir de formes 3D.

Mots-clés Biométrie Douce, Visage 3D, Classification du Genre, Estimation de l'Age, Géométrie Riemannienne.

Abstract Soft-Biometric (gender, age, etc.) recognition has shown growing applications in different domains. Previous 2D face based studies are sensitive to illumination and pose changes, and insufficient to represent the facial morphology. To overcome these problems, this thesis employs the 3D face in Soft-Biometric recognition. Based on a Riemannian shape analysis of facial radial curves, four types of Dense Scalar Field (DSF) features are proposed, which represent the Averageness, the Symmetry, the global Spatiality and the local Gradient of 3D face. Experiments with Random Forest on the 3D FRGCv2 dataset demonstrate the effectiveness of the proposed features in Soft-Biometric recognition. Furtherly, we demonstrate the correlations of Soft-Biometrics are useful in the recognition. To the best of our knowledge, this is the first work which studies age estimation, and the correlations of Soft-Biometrics, using 3D face.

Keywords Soft-Biometrics, 3D face, Gender Classification, Age Estimation, Riemannian Geometry.