Joint Gender, Ethnicity and Age Estimation from 3D Faces An Experimental Illustration of their Correlations

Baiqiang Xia, Boulbaba Ben Amor, Mohamed Daoudi

Noname manuscript No.
(will be inserted by the editor)

Joint Gender, Ethnicity and Age Estimation from 3D Faces

An Experimental Illustration of their Correlations.

Baiqiang Xia · Boulbaba Ben Amor · Mohamed Daoudi

Received: date / Accepted: date

Abstract Humans present clear demographic traits which allow their peers to recognize their gender and ethnic groups as well as estimate their age. Abundant literature has investigated the problem of automated gender, ethnicity and age recognition from facial images. However, despite the coexistence of these traits, most of the studies have addressed them separately, very little attention has been given to their correlations. In this work, we address the problem of joint demographic estimation and investigate the correlation through the morphological differences in 3D facial shapes. To this end, a set of facial features are extracted to capture the 3D shape differences among the demographic groups. Then, a correlation-based feature selection is applied to highlight salient features and remove redundancy. These features are later fed to Random Forest for gender and ethnicity classification, and age estimation. Extensive experiments conducted on FRGCv2 dataset, under Expression-Dependent and Expression-Independent settings, demonstrate the effectiveness of the proposed approaches for the three traits, and also show the accuracy improvement when considering their correlations. To the best of our knowledge, this is the first study exploring the correlations of these facial soft-biometric traits using 3D faces. This is also the first work which studies the problem of age estimation from 3D Faces.[1]

Keywords 3D Face · Gender · Ethnicity · Age · Correlation · Random Forest · Feature Selection.

1 Introduction

In daily life, human beings perform gender and ethnicity recognition as well as estimate the age of their peers naturally and effectively. Several studies from different backgrounds (face and head anthropometry, cognitive psychology, clinical studies, etc.) have tried to understand how the process works. In particular, a number of anthropometric studies [65] have revealed that significant facial morphology differences exist among the gender, the ethnicity and the age groups. For example, when studying the *Sexual Dimorphism* (Male/Female differences) [9], 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. [65] have also demonstrated that all the concerned anthropometric measurements of females are smaller. In the study of the ethnic differences [16], researchers have found that compared to the North America Whites, Asians usually have broader faces and noses, far apart eyes, and exhibit the greatest difference in the anatomical orbital regions (around the eyes and the eyebrows). In the clinical study reported in [35], *Alphonse et al.* have revealed that Caucasians have significantly lower fetal Fronto-Maxillary Facial Angle (FMFA) measurements than Asians. In [65], sixteen anthropometric measurements have been recognized as significantly different between Asian and Caucasian faces. When studying the face aging [48, 49], researchers have concluded that the craniofacial growth is the main change in baby and adolescent faces, which results in the re-sizing and redistribution of facial features. During this period, generally, the larger is the age, the bigger is the size of the face. 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 a triangle shaped face with small amount of wrinkles. In contrast, old adults are usually associated with a U-shaped face with significant wrinkles on the face. Besides the existence of these *Soft-Biometric Traits*[2] [11, 34] in the face, gender, ethnicity and age are also correlated in characterizing the facial shape [65]. For example, according to the anthropometric studies cited above, the shape of the nose is influenced by all the three soft-biometric traits. In human perception, female faces usually look smoother and younger than male faces, and the Asian faces usually look younger than Non-Asian faces [50]. In [58], *Vignali et al.* have

Baiqiang Xia, Boulbaba Ben Amor, Mohamed Daoudi
IMT Lille Douai, Univ. Lille, CNRS, UMR 9189 - CRIStAL-Centre de Recherche en Informatique Signal et Automatique de Lille, F-59000 Lille, France. E-mail: {xia.baiqiang; boulbaba.benamor; mohamed.daoudi}@telecom-lille.fr

[1] Part of this work has been published in the International Conference on Computer Vision Theory and Applications 2014 [61] and won the **Best Paper Award** in the area of **Image and Video Understanding**.

[2] A.K. Jain defined Soft-biometric traits as a set of traits providing information about an individual, though these traits are not sufficient to individually authenticate the subject because they lack in distinctiveness and permanence [34].

demonstrated both visually and quantitatively that ethnicity and gender are correlated to some extend in the 3D face. In [19], *Gao et al.* have concluded that the gender classifier trained on a specific ethnicity could not get good generalization ability on other ethnic groups. In [58], when the gender information is removed from the faces, the human ethnicity classification performance is recognizably lower.

In this paper, we consider the problem of joint gender, ethnicity and age recognition through the morphological differences in the 3D facial shapes. A number of morphology-driven 3D features are extracted, and then used to train Random Forest in Classification/Regression modes for gender and ethnicity recognition, and age estimation. A quantitative study of the correlations of the three demographic traits is proposed. In the joint recognition experiments, these correlations are considered, to demonstrate their influences on the recognition accuracy.

1.1 Related Work

Since the 90s of the previous century, several approaches have been proposed to solve the problem of image-based automatic facial Soft-Biometric Traits recognition, e.g. gender [22,26,45], ethnicity [17,26] and age [18,22,26,48].While the most conventional works have attempted to exploit the intensity (color) images, a recent trend consists on investigating the use of the 3D shape of the face. According to the definition given in [62], the face texture represents the reflection and absorption effects of external illumination caused by the facial skin, while the 3D face shape defines the solid border which distinguishes the face and the environment. It is now well-established that the 3D shape provides a rich description of the face morphology compared to the intensity image. From 2D intensity images, the (2D) shape of the face is usually represented by a sparse set of anchor points detected in the face images (used to define the well-known Active Shape and Active Appearance Models), which represent poorly and incompletely the facial shape, and are sensitive to the head pose changes during the image acquisition. In the study presented in [29], *Hu et al.* have demonstrated that, with the 3D shapes of the face, human observers perform better on both gender and ethnicity recognition than with the 2D images.

With a particular focus on the 3D shape-based methods, this paragraph provides a brief review of the existing work in gender, ethnicity and age recognition. We adopt the taxonomy proposed recently in [26] which consists of three categories: (1) anthropometry-based, (2) Image-based, and (3) Appearance-based in automatic demographic traits recognition. The recent work of *Gilani et al.* [21] belongs to the first category (anthropometry) as it has proposed to automatically detect the biologically significant 3D facial landmarks, and then calculate the Euclidean and the Geodesic (along the surface) distances between them as face features. Similar studies have been proposed in [58] and [27] where the 3D facial landmarks coordinates [58], and the volume/area information of facial regions [27] are extracted for gender and ethnicity classification. Despite their high performance, these approaches require accurate detection of the anchor points on the face. The works of *Toderici et al.* [54], and *Wang and Kambhamettu* [59] belong to the second category where image-based features like the wavelets [54], the Local Binary Patterns (LBP) [59], and the Shape Index [59] are extracted from the range images for gender and ethnicity recognition. With 3D facial meshes, *Tokola et al.* [55] have extracted the Correspondence Vectors (CVs) features to recognize all the three facial traits. Approaches which combine 2D (texture) and 3D (shape) channels form the third category (appearance-based). In [30], *Huang et al.* have combined the Local Circular Patterns (LCP) features of texture and range images in gender and ethnicity recognition. Also in [60], *Wu et al.* have fused shape and texture information implicitly with needle maps recovered from intensity images. In [33], *Huynh et al.* have fused the Gradient-LBP from range images and the Uniform LBP features of the gray image in gender classification. In [44], *Lu et al.* have fused the SVM outputs from range and intensity images in gender and ethnicity classification. Reported results of the appearance-based demonstrate higher accuracy than using only the intensity or the depth. However, when using the 2D intensity or range images, manually labeled facial landmarks are usually required in a pre-processing step to crop and align the facial regions [26,30,44].

From the analysis above, most of the existing 3D-based methods have tried to extract conventional features from range images (such as the 3D landmark coordinates, LBP, LCP, wavelets, shape index, etc.). No attention has been paid to reveal the relationship between the extracted features and the studied demographic trait. That's to say, even an approach achieves high recognition performance, we don't know why the extracted features are relevant in the studied task. Moreover, the study of these soft-biometric traits have been done separately, in the 3D domain. The correlations of the soft-biometric traits have attracted little consideration. Although some 2D-based works have investigated the relationship between ethnicity and gender [15,19], the relationship between ethnicity and age [24,38], and the relationship between gender and age [23, 24, 38, 46, 48, 56] in their recognition tasks, different conclusions have been reported. For example, experimental results in [15] and [19] have made different answers to the question whether gender and ethnicity are helpful in each others' recognition. The various illumination conditions and head poses in different data acquisitions, the dependency on the accuracy of the landmarks in face alignment, and especially the incomplete facial shape information in the 2D images, should account largely for this disagreement. We propose here to explore the morphological differences in the 3D shape of the face to answer the following two questions: **(1)** Can the 3D shape of the face reveal our gender, ethnicity and age? and **(2)** are the correlations between the demographic traits useful in each others' recognition task? To the best of our knowledge, this is the first work in literature that proposes a joint estimation of the demographic traits through the 3D shape differences. It is also the first work which studies the problem of age estimation from 3D faces [61].

1.2 Methodology and Contributions

Our approach consists of a 3D feature extraction step followed by a classification or regression step. The 3D face scans are first pre-processed to extract the facial region, then a collection of radial curves emanating from the nose tip are extracted to represent the face. Following four different pairwise curve comparison strategies, we compute four types of 3D features, for which each reflects a specific perspective in face perception. In the classification/regression step, we present the features to Random Forest in its classification mode for gender or ethnicity classification, and in regression mode for age estimation. To enhance the performance, we have also included two additional steps in our approach, the *Feature Selection* and the *Fusion*. The feature selection method is used for selecting a salient subset of features which contains the information of gender, ethnicity and age. The fusion method merges all the information from the four features by concatenation. The main contributions of this work are the following:

- We propose a set of 3D facial descriptors grounding in Shape Analysis of facial radial curves, with which we demonstrate that **3D shapes of our faces can reveal our gender, ethnicity and age information**. These descriptions are designed to capture in different ways the morphological difference among the demographic groups.
- We demonstrate that **gender, ethnicity and age are correlated in the 3D face**, and their correlations are helpful in each others' recognition. Our conclusion is significantly different than [15], which claims that gender and ethnicity are not helpful for each others recognition.
- This is the first work in the literature which investigates the problem of **age estimation from 3D shapes** and **perform joint estimation of gender, ethnicity and age in 3D**. Extensive evaluations on the challenging FRGC dataset demonstrate the effectiveness of the proposed method and its robustness to facial expression variations.

The rest of the paper is organized as follows. In Section 2, the computational strategies of the 3D facial morphological features are detailed, as well as their relationship with the demographic groups. Section 3 explains the machine learning techniques adopted in our work for facial soft-biometric traits recognition. Experiments and discussions are issued in Section 4. In the end, Section 5 makes the conclusions and states some future directions.

2 Geometrical 3D Features Extraction

In this section, we describe four different and complementary morphological face descriptions extracted from the 3D face. These descriptions are densely computed based on shape analysis of 3D radial curves of the face. Earlier studies on 3D face recognition [13] and 4D expression recognition [7] have demonstrated the effectiveness of the proposed geometrical framework in comparing 3D faces and capturing shape deformations. We point out that although in the common geometrical shape analysis background with [7], this work presents several methodological and applicational contributions, which will be highlighted later.

2.1 Mathematical Background

Following the previous studies [7,13] in 3D face analysis, the 3D scans are first preprocessed by filling the holes, cropping the informative (face) area and denoising the 3D mesh, together with nose tip detection and pose normalization. Then, a collection of radial curves stemming from the nose tip are defined over the face. The latter step 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 radial curves, respectively. To allow appropriate shape analysis of facial curves, we adopt the mathematical representation proposed in [51], termed *Square Root Velocity Function* (SRVF), **which has the merit of allowing the shape registration and the shape analysis of elastic curves simply with the same $\mathbb{L}^2$ metric**.

Formally, 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* (SRVF), q, according to :

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

Then, with the $\mathbb{L}^2$-norm of the q functions scaled to 1 ($\|q\| = 1$), the space of such representation: $\mathcal{C} = \{q : [0,1] \to \mathbb{R}^3, \|q\| = 1\}$ becomes a Riemannian manifold and have a spherical structure in the Hilbert space $\mathbb{L}^2([0,1],\mathbb{R}^3)$. In virtue of the spherical structure of $\mathcal{C}$, illustrated in Fig. 1, given two curves β_1 and β_2 represented by their SRVFs q_1 and q_2 on the manifold, the geodesic path connecting q_1, q_2 is given analytically by the minor arc of the great circle connecting them on $\mathcal{C}$ (see [51] for further details). Our goal here is to effectively quantify the shape divergence between β_1 and β_2, where an accurate registration is required to match similar morphological parts between the curves. For example, to capture the bilateral symmetry on a pair of symmetrical curves in a face, it is important to accurately match them so that the peaks match the peaks and valleys go with the valleys. It has been proofed in [51] that under the $\mathbb{L}^2$-norm, the quantities $\|q_1 - q_2\|$ and $\|q_1 \circ \gamma - q_2 \circ \gamma\|$ are same, where the composition $(q \circ \gamma)$ denotes the function q with a new parameterization dictated by a non-linear function $\gamma : [0,1] \to [0,1]$. This important property allows curves registration by re-parameterization, and thus makes the curves registration easier. In fact, it allows to consider one of the curves as reference and search for a $\gamma^* = \text{argmin}_{\gamma\in\Gamma}(\|q_1 - \sqrt{\dot{\gamma}} q_2 \circ \gamma\|)$ which optimally registers the two curves. This optimization is resolved by Dynamic Programming, as described in [51] for general $\mathbb{R}^n$ curves.

After the registration (let q_2 becomes q_2^* with the optimal re-parameterization function γ^*), we need to quantify the shape difference of the two curves. Due to the spherical structure of $\mathcal{C}$, the geodesic path connecting two points q_1 and q_2^* has a constant velocity. Thus, we use the tangent vector $V_{q_1 \to q_2^*}$, element of the tangent space on q_1 to the manifold $\mathcal{C}$, $T_{q_1}(\mathcal{C})$, to quantify the shape difference from q_1 to q_2^*. This vector is also the initial shooting vector along the geodesic connecting q_1 towards q_2^*. Again, due to the spherical structure of $\mathcal{C}$, $V_{q_1 \to q_{2*}}$ is easy to compute using the inverse exponential map given by:

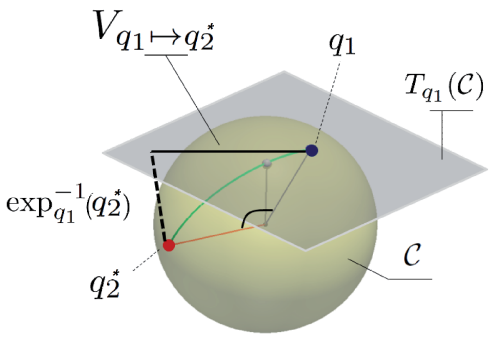


Fig. 1 An illustration of the spherical structure of the manifold $\mathcal{C}$, q_1 (blue) and q_2^* (red) are elements of $\mathcal{C}$. The geodesic connecting them (green path), $T_{q_1}(\mathcal{C})$ (gray) is the tangent space of $\mathcal{C}$ on q_1, and the shooting vector $V_{q_1 \to q_2^*}$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

$$V_{q_1 \to q_2} = \exp_{q_1}^{-1}(q_2^*) = \frac{\theta}{\sin(\theta)}(q_2^* - \cos(\theta) q_1),$$

Where q_2^* is q_2 re-parameterized by the optimal matching function γ^* that provides the best registration of q_2 with respect of q_1, and $\theta = \cos^{-1}\langle q_1, q_2^* \rangle$ is the length of the geodesic connecting q_1 to q_2^*.

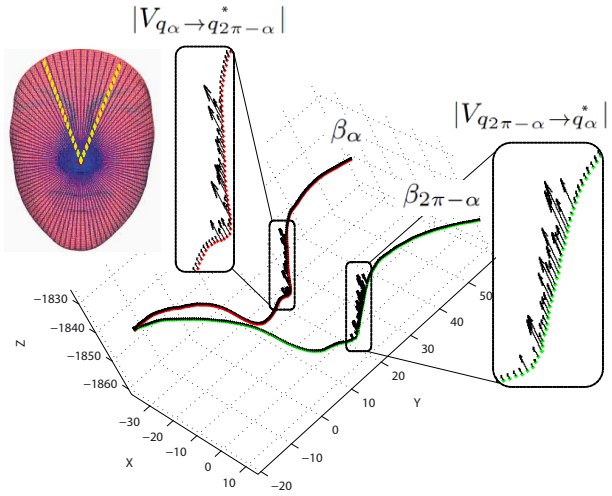


Fig. 2 Illustration of the deformations needed to fit β_α to $\beta_{2\pi-\alpha}$ ($|V_{q_\alpha \to q^*_{2\pi-\alpha}}|$) and inversely ($|V_{q_{2\pi-\alpha} \to q^*_\alpha}|$), on each vertex of the curve parameterized by t. The quantities q_α and $q_{2\pi-\alpha}$ denote the SRVFs of the curves β_α and $\beta_{2\pi-\alpha}$, respectively, extracted from a given face.

An illustration of $|V_{q_1(t) \to q_2^*(t)}|$ between points of two symmetrical curves on the same face is shown in Fig. 2. Here $|.|$ denotes the magnitude of $V_{q_1(t) \to q_2^*(t)}$ on each parameter t. The arrows shown on the facial curves capture the point-wise local deformations needed to go from the reference curve β_1 to the target (symmetrical) curve β_2^* (or from reference curve β_2 to target curve β_1^*). We note that, in general, the quantities $|V_{q_1(t) \to q_2^*(t)}|$ and $|V_{q_2(t) \to q_1^*(t)}|$ are different, due to the re-parameterization on the target curve in the registration step.

2.2 Morphology-driven 3D Features Computation

Based on the mathematical framework described above, we extract four different facial descriptions, to capture the morphological differences among different gender, ethnicity and age groups. Each of them reflects a perceptible property in the human face, namely the bilateral **Symmetry**, the face **Averageness**, the global (**Spatial**) variations and the local variations (**Gradient**). These descriptions are illustrated in Fig. 3, where in each panel, the left part illustrates the extracted radial curves and the curve comparison strategy, and the right part shows the extracted features as an interpolated color-map on the face. On each vertex of the mesh, the colder is the color, the higher is the magnitude of the shape difference.

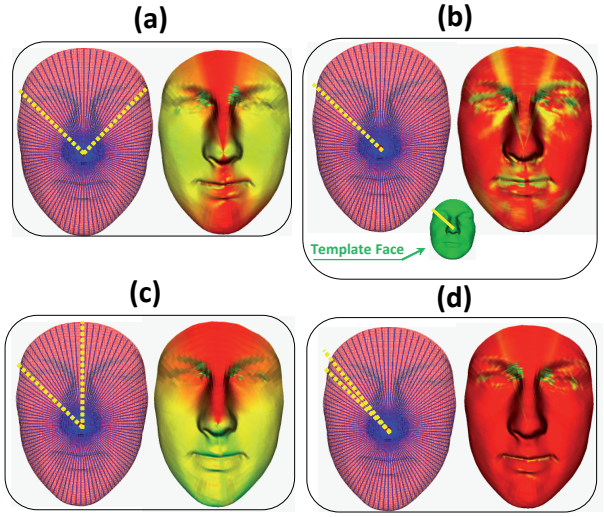


Fig. 3 Illustrations of different features on 3D shape of the face $\boldsymbol{S}$. (a) 3D-Symmetry Description: shape differences from each radial curve $\beta_\alpha^{\boldsymbol{S}}$ to its symmetrical curve $\beta_{2\pi-\alpha}^{\boldsymbol{S}}$; (b)3D-Averageness Description: shape differences from radial curve $\beta_\alpha^{\boldsymbol{S}}$ in a preprocessed face to radial curve $\beta_\alpha^{\boldsymbol{T}}$ in face template (with same angle index α); (c) 3D-Spatial Description: shape differences from radial curve $\beta_\alpha^{\boldsymbol{S}}$ to the middle-up radial curve $\beta_0^{\boldsymbol{S}}$ in the forehead; (d) 3D-Gradient Description: shape differences from radial curve $\beta_\alpha^{\boldsymbol{S}}$ to its neighbor curve $\beta_{\alpha+\Delta\alpha}^{\boldsymbol{S}}$.

– **The 3D-Symmetry Description (3D-sym.)** shown in Fig. 3(a) captures densely the deformation between bilateral symmetrical curves ($\beta_\alpha^{\boldsymbol{S}}$ and $\beta_{2\pi-\alpha}^{\boldsymbol{S}}$). The colormap shown on the right illustrates this description on each vertex of the 3D face. Here colder colors stand for higher feature values. In other words, they highlight the most asymmetric areas in the 3D shape of the face. This description reflects the idea that different population has different bilateral face symmetry. In [42], *Little et al.* have revealed that the facial symmetry and sexual dimorphism are related in humans, and faces of different ethnicity have different amount of facial asymmetry. In [52], *Steven et al.* have found the facial masculinization significantly covaries with the fluctuating asymmetry in men's face. In [10], *Clinton et al.* have discovered that increasing age is associated with a higher degree of facial asymmetry in 3D face surfaces.

– **The 3D-Averageness Description (3D-avg.)** shown in Fig. 3(b) compares a pair of curves with the same angle index α from a face $\boldsymbol{S}$ and an average face template $\boldsymbol{T}$. The average face template $\boldsymbol{T}$ (presented in Fig. 3(b)) is constructed as the middle point of geodesic path connecting a representative male face to a representative female face. This description reflects the idea that different population shows different shape size and anthropometric features. Researchers have revealed that facial sexual dimorphism relates closely to facial distinctiveness (the converse to averageness) [6]. The male faces usually possess more prominent features than female faces [32, 57, 65]. For ethnicity, the Asian and Non-Asian population present significantly different facial morphology [2, 14, 40, 65], such as the face width, and the nose width and height. Also, in [1, 12], researchers have found that exaggerating distinctiveness in the face produces an increase in the apparent of its age.

– **The 3D-Spatial Description (3D-spat.)** shown in Fig. 3(c) captures the shape divergence 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 $\boldsymbol{S}$. As the curve β_0 is the most rigid curve in the face, the 3D-Spatial description aims to capture the smoothly-changing global deformation from the most rigid part of the face. As demonstrated in [36, 64], sexual dimorphism exhibits unequal magnitude in different facial parts. In [3], *Ashok et al.* have found that the facial features contribute more than the nose and head features towards sexual dimorphism. For ethnicity, Asian and Non-Asian faces exhibit different morphological differences in different facial parts. For age, it's well-known that different facial parts age differently [39,53]. The internal facial parts (especially the eyes surrounding areas) are recognized as the most significant for automatic age estimation [39].

– **The 3D-Gradient Description (3D-grad.)** shown in Fig. 3(d) computes the differences between pairwise neighboring curves ($\beta^{\boldsymbol{S}}_\alpha$ and $\beta^{\boldsymbol{S}}_{\alpha+\Delta\alpha}$). It takes the idea that different population could show different local shape changes in the face. For gender, it has been revealed that sexual dimorphism demonstrates the facial developmental stability [41] in the face. For ethnicity, the Asians usually have wider face and nose, and less protruding nose, than the Non-Asian faces. For age, this description reflects the local facial changes (loose of strength and smoothness when aging) which are very important in age perception [28,41].

It is clear that the 3D descriptors described above reflect (quantify) different cues of the facial morphology. They allow to capture densely differences which exist across the studied demographic groups. The question now is – **Can these 3D morphological descriptions reveal our demographic traits?**

3 Facial Soft-biometric Traits Recognition

We consider now the remaining steps in our facial demographic traits recognition algorithms. First, we perform a correlation-based Feature Selection method on our descriptions. Then, we present the selected features to Random Forest for facial classification/regression. We note that these two steps are common in many recognition problems, however, we shall present here in-deep analysis on the correlation between the facial demographic traits and the extracted 3D features. In particular, we shall highlight the location of the salient features on the face for age, gender and ethnicity recognition.

3.1 Correlation-based Feature Selection

Feature subset selection is the process of identifying and removing as much irrelevant and redundant information as possible [25]. 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 [37]. We choose a filter method for feature selection, named the Correlation-based-Feature-Selection [25], 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 attributes, the gender (labeled as male and female), the ethnicity (labeled as Asian and Non-Asian) and the age (labeled in two groups, > 22 group and < 23 group). After Feature selection, we retain $200 - 400$ features for each description. Thus, the feature selection procedure significantly reduces the size of the features. To illustrate this idea and allow analysis of the location of salient facial areas for each demographic trait, in Figure 4, we show the magnitude of the correlation between the facial features and the facial attributes as color map on the face. The correlations are calculated with the 466 earliest scans of FRGCv2 dataset as follows.

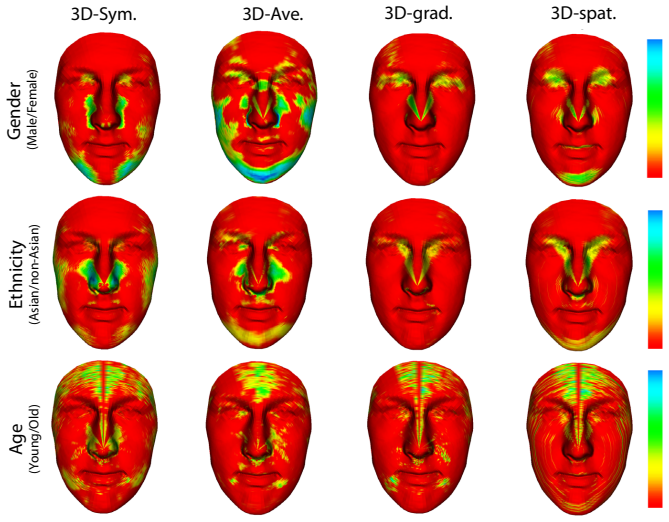


Fig. 4 Correlations between 3D facial features and demographic traits on each vertex of the face (colder colors indicate higher correlations).

For all the earliest (466) scans of FRGCv2 dataset, the *ith* feature of a face description makes a one dimensional vector $D_i = (d_i^1, d_i^2, ..., d_i^{466})$, where d_i^m denotes the *ith* feature of the *mth* face. Then, for a given facial demographic trait, the labels for the same 466 scans make a vector $L_a = (L_1, L_2, ..., L_{466})$, where L_m denotes the attribute label of the *mth* face. 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. Recall that, for two variables X and Y, the *Pearson correlation coefficient* $p_{X,Y}$ is defined by $p_{X,Y}$=**cov**$(X,Y)/(\sigma_X \sigma_Y)$, where **cov** denotes the covariance and σ denotes the standard deviation. The absolute value of the *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, the more informative is this feature for discriminating the concerned demographic trait. In Figure 4, the correlation between each description and each facial attribute is shown as color-map on each vertex of the face. Warm colors indicate low correlations, and cold colors indicate high correlation. From the figure one can note several observations.

Firstly, the green/blue colors show that the facial features are considerably correlated with the facial soft-biometric traits, in some particular areas of the face. For gender, the eyes, nose, cheek-sides, lips and the chin are particularly informative. It matches with the previous findings of sexual dimorphism presented in [9], 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 echoes the findings in [16] which have stated that the Non-Asians have broader faces and noses, farther apart eyes, and lower fetal fronto-maxillary 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 the forehead, eyes, nose and mouth regions with age.

Secondly, it confirms that the four descriptions give different and complementary perspectives for perception of gender, ethnicity and age. They act as four different types of 'eyes' in face perception. For gender (first row), the *3D-sym* description 'sees' the inner eye corners, the border area of the nose and the cheeks, and the chin-side regions. The *3D-ave* description 'looks' at the eyebrows, the eyes, the nose, the lips, and gives significant attention to 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 emphasizes the chin and the sides of the nostrils. For ethnicity (second row), the *3D-sym* description 'sees' the nose regions and the cheeks. The *3D-ave* description 'looks' at the border 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 (third row), the *3D-sym* description 'sees' the whole forehead, the nose, the outer eye corners, and 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 central part of the forehead, the regions around the eye corners, and the nose bridge. According to these observations, the four descriptions capture different and complementary cues of gender, ethnicity and age in the face.

3.2 Random Forest Classification/Regression

Random Forest is an ensemble learning method that grows many decision trees $t \in \{t_1, .., t_T\}$ considering an attribute [8]. 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 $\boldsymbol{N}$ instances are sampled randomly with replacement from the original data. Then at each node of the tree, a constant number of $\boldsymbol{m}$ $\boldsymbol{(m<<M)}$ variables are randomly selected, and the best split on these $\boldsymbol{m}$ 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 $\boldsymbol{m}$ reduces both the correlation and the strength. Increasing it increases both. Thus, an optimal $\boldsymbol{m}$ is needed for the trade-off between the correlation and the strength. In Random Forest, the optimal value of $\boldsymbol{m}$ 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 Experimental Results

In the following experiments, we use the Random Forest method [8] in classification mode for gender and ethnicity classification, and the regression mode for age estimation. For the experiments in section 4.1-4.3, they are carried out on the Face Recognition Grand Challenge 2.0 (FRGCv2) dataset [47]. The FRGCv2 dataset was collected by researchers from the University of Notre Dame and contains $4,007$ 3D face scans of 466 subjects with differences in gender, ethnicity, age and expression. For gender, there are $1,848$ scans of 203 female subjects and $2,159$ 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 $2,554$ scans of 319 White subjects, $1,121$ 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. With this dataset, we conduct two types of experiments as follow:

- **Expression-Dependent Experiment** uses the 466 earliest scans 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. [5] and [21] have explored this data subset for 3D gender classification.
- **Expression-Independent Experiment** is based on the whole 4,007 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 [13]. In [63], the ethnicity classification results on FRGCv2 dataset are influenced strongly by the facial expressions.

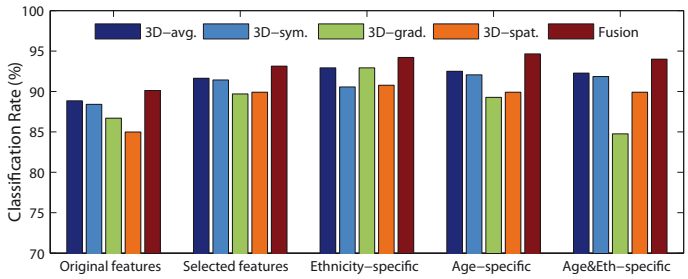
We use the *Leave-One-Person-Out (LOPO)* cross-validation approach in these experiments, where each time the scans of one subject are used for testing, and the scans of the remaining subjects are used for training. Thus, there are altogether 466 folds in the cross-validation. The experiments are conducted in a Subject-independent fashion. Each subject is tested only once. The LOPO strategy is similar to real-word like applications, and it allows training with a maximum number of subjects. To give consideration to the correlations of the three attributes, we define the *Attribute-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, then we evaluate 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, then test on each ethnicity group separately. For *Age-specific* setting, we separate the 466 subjects into older group ($\geq$26 years, 107 subjects) and younger group ($\leq$25 years, 359 subjects) first, then perform LOPO experiments on the younger and older groups, separately. Note that in section 4.5, we will use a different experimental setting which consists on 10-fold cross-validation (10-CV) to allow comparison to previously published approaches.

4.1 Gender Classification on FRGCv2

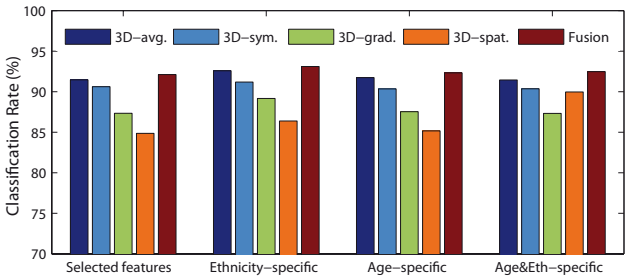
In this experiment, we first perform LOPO gender classification with Random Forest on the original extracted features. Then, we perform correlation-based feature selection on the original features and carry out the experiments with the selected features. After that, we perform the experiments in consideration of ethnicity and age in the *Ethnicity-specific* and *Age-specific* settings, respectively. We consider also *Eth&Age-specific* experiment, where the experiment is conducted with the scans of the same ethnicity and age groups.

The gender classification results are shown as bar-plots in Fig. 5. The y-axis gives the classification rate in LOPO experiment. The x-axis shows the different experimental settings. The left panel corresponds to the Expression-Dependent experiments, the right one shows the results when tolerate facial expression variations (Expression-Independent experiments). With the 466 scans in the Expression-Dependent experiments, the original features achieve more than 85% gender classification rate for each description. With the feature selection step, the results are improved by $2-5\%$. Now, when considering the ethnicity information in the *Ethnicity-specific* setting, the results are furthermore improved in general. **The improved results demonstrate that Asian and Non-Asian exhibit different gender information**. When considering the age information in the *Age-specific* setting, a stronger improvement is shown, **which indicates that people of different age exhibit different gender information**. When considering both *Ethnicity-specific* and *Age-specific*, termed *age&Eth-specific*, the accuracy is generally higher than the Feature Selection and quite comparable to *Ethnicity-specific* and *Age-specific*. We also note that the fusion of the 3D features always outperforms individual features in all the settings. The highest gender classification rates, 94.64% and 94.21%, are achieved by the *Fusion* under the *Age-specific* and *Ethnicity-specific* settings, respectively. These findings are furthermore confirmed by in the Expression-Independent experiments, as shown in the right panel of Fig.5. 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 performance. Among the four descriptions, generally, the *3D-avg* descriptor, which captures the shape difference to a given template, achieves the highest performance. It confirms the studies on sexual dimorphism [9] which claim that Male and Female faces present different morphological features in the shape. Also, the *3D-sym* descriptor confirms that facial asymmetry is related to the gender [43]. In addition to these confirmations, these experiments show the relationship between the gender information and two influencing intrinsic factors which are the age and the ethnicity. However, concerns should be raised as the gender distributions may not be uniform across the demographic subgroups of FRGCv2, which could cast different prior information in to the classifiers. To this end, we applied the Two-sample t-test (with null hypothesis of equal Mean in two samples), and the Two-sample F-test (with null hypothesis of equal Variances in two samples) to the subjects' gender in different demographic subgroups. Considering different ethnicity subgroups, the tests results (t-test: $p = 0.1752$, F-test: $p = 0.8090$) suggest there is no significant difference of gender distributions between the Asian and Non-Asian subjects in terms of Mean and Variance, at a significant level of 0.05. For age, results (t-test: $p = 0.1565$, F-test: $p = 0.8061$) demonstrate again that there is no significant difference between the gender distribution in the Old (>=26) and the Young

Fig. 5 gender classification results under Expression-Dependent (ED) and Expression-Independent (EI) 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.

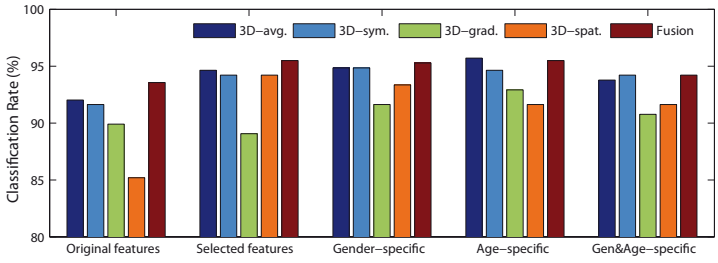


(a) Expression-dependant gender classification results

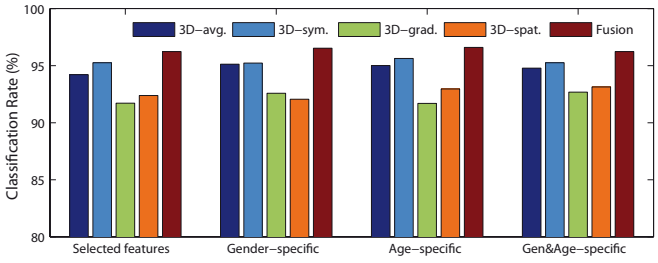


(b) Expression-independant gender classification results

Fig. 6 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.



(a) Expression-dependant ethnicity classification results



(b) Expression-independant ethnicity classification results

(<=25) subgroups. These statistical tests results reveal that the gender classification improvements in *attribute-specific* settings are unlikely to be linked to the gender distribution differences in the subgroups. In other words, the improvements are related to the *attribute-specific* setting itself which highlights specific gender features in different demographic subgroups.

4.2 Ethnicity Classification on FRGCv2

Ethnicity classification consists on automatically label a query instance into its corresponding ethnicity class (Asian or Non-Asian, in the present study). Similarly to the experiments conducted in section 4.1 for gender classification, the LOPO experiments are conducted for ethnicity classification under the Expression-Dependent and Expression-Independent settings. We have also explored the use of the original features, the selected features, and then study in *Gender-specific* and *Age-specific* settings. The age partition is the same as in the *Age-specific* setting for gender classification. The ethnicity classification results are shown in Fig.6. As shown in the left panel, under the Expression-Dependent setting, the results from the original features are always higher than 85%. The feature selection improves the results with $2\% - 7\%$ compared to the previous results. Under the *Gender-specific* and *Age-specific* settings, the results are slightly higher. **The enhancements in results demonstrate that the Male and the Female have different ethnicity information, and people of different age have different ethnicity information**. Again, the highest ethnicity classification rates of 95.71% and 95.49% are achieved by the *3D-avg* description and the fusion, respectively. Also, the fusion of these features almost always outperforms each individual description in each setting. These results are confirmed in the right panel of Fig. 6 with a higher accuracy of 96.6%. This demonstrates the robustness of the proposed approach against the facial expressions in ethnicity classification. According to these results, ethnicity (Asian and Non-Asian) classification is influenced by the gender and age factors. The *3D-avg.* description achieves higher accuracy, compared to the remaining descriptions. It confirms the findings of previous studies [16,35] that a significant difference exists between Asian and Non-Asian faces. In addition, according to the results in Fig.6, the bilateral asymmetry *(3D-sym)* can play an important role in ethnicity classification. Similar to gender study, we applied the Two-sample t-test, and the Two-sample F-test to the subjects' ethnicity in different demographic subgroups, to reveal the significance of distributional differences.

Considering gender subgroups, the tests results (t-test: p = 0.1752, F-test: p = 0.2459) suggest there is no significant difference of ethnicity distributions between the Male and Female subjects in terms of Mean and Variance. These results demonstrate that the observed improvement in *gender-specific* ethnicity recognition is unlikely to be linked to the distributional difference of gender information in Asian and Non-Asian subjects. However, for age subgroups, tests results (t-test: p = 7.764e-23, F-test: p = 2.637e-08) demonstrate that ethnicity distributions in the Old (>=26) and the Young (<=25) subgroups are of significant difference. It suggests a confounding factor exists in *age-specific* ethnicity recognition on FRGCv2, as the distributional difference could project a more accurate prior information of ethnicity during the training of classifier in this specific case. While, due to unavailability of more proper 3D face dataset, the effect of this factor could not be studied in the current stage.

4.3 Age Estimation on FRGCv2

Given a face image, estimating its age consists on automatically label it with an exact age or an age group/range. In the following, we use the Random Forest method in regression mode to estimate the exact age of a query. Similar to the previous two experiments for gender and ethnicity classifications, we compare the results achieved with the original features, results from the selected features, and results reported under *Gender-specific* and *Ethnicity-specific* settings. We note here, in feature selection, we use a **different** age group partition than in the *Age-specific* experiments. The 466 subjects are divided into two age groups, one > 22 years which represent 162 subjects, and another group < 23 years which includes 304 subjects. The idea behind this partition is the fact that the craniofacial growth stops at the age of $18-20$, and faces exhibit different aging morphology before and after this age. In addition, it balances better the number of scans in the two groups. 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 in [20] to evaluate age estimation algorithms, shows the percentage of cases among the testing set where the absolute age estimation error is less than a threshold. In this work, the experimental results are shown as MAEs, in Fig. 7.

As shown in the left panel of Fig. 7, under the Expression-Dependent experiments, the MAEs for all the descriptions are always under 4 years. After feature selection, the errors decrease which confirms again the usefulness of feature selection to highlight the salient features in our approach. When testing under the *Ethnicity-specific* and the *Gender-specific* settings, the errors decrease again. When considering both gender and ethnicity (*Gen&Eth* setting), the MAEs are even lower. It demonstrates that **the knowledge of gender and ethnicity can improve age estimation accuracy on the FRGCv2 dataset, which suggests that different gender and ethnicity could possess different facial aging patterns.** The **Fusion** provides the highest performance in each setting. The lowest MAE is achieved in the *Gen&Eth-specific* setting, by the **Fusion** with 2.95 years. These observations are confirmed by the results of the Expression-Independent experiments, as depicted in the right panel of Fig 7. The *Fusion* always outperforms individual descriptions, and achieves the lowest MAE of 3.24 years in the *Gen&Eth-specific* setting. Because of the novelty of this problem of 3D age estimation, we shall provide later, in Section 4.5, a comparative study of the proposed work (using 3D) and the recent published work [31] (based on 2D images).

Table 1 MAEs achieved by the fusion for different age groups under Expression-Dependent (ED) and -Independent (EI) settings.

Fusion/age groups	≤ 20	(20,30]	(30,40]	> 40	Overall
ED Experiment					
Original features	3.93	2.29	7.03	24.45	3.63
Selected features	2.88	2.18	8.80	23.26	3.42
Gender-specific	2.79	1.93	8.64	19.84	3.14
Eth-specific	2.62	2.06	8.37	22.27	3.20
Gen&Eth-specific	**2.48**	**1.86**	**8.12**	**19.69**	**2.95**
EI Experiment					
Selected features	2.92	2.22	8.15	24.03	3.75
Gender-specific	2.57	2.16	7.76	23.39	3.55
Eth-specific	2.71	2.11	7.36	20.67	3.43
Gen&Eth-specific	**2.36**	**2.03**	**7.39**	**20.01**	**3.24**
# of Subjects	**185**	**246**	**20**	**15**	**466**

We provide in Table 1 the age estimation accuracy related to each age group. We note that in both Expression-Dependent and Expression-Independent experiments, the MAEs in an age group is always lower when considering gender and ethnicity information. When considering both gender and ethnicity in age estimation, the MAEs in each age group always reach the lowest values (marked in bold). Thus, by giving consideration to gender and ethnicity, we have successfully enhanced the age estimation performance for all the age groups. Table 1 also shows that our algorithm performs much better in young age groups, than in old age groups. Considering the number of training subjects, as shown in the last row of the table, it is probably due to the unbalanced distribution of the scans across the age groups. A limited number of scans are available for older age groups.

The age estimation results on FRGCv2 demonstrate that, knowing the gender and ethnicity information, the improvement of performance is significant. However, like for gender and ethnicity recognition, concerns should be raised on the age distribution differences in different demographic subgroups. With the Two-sample t-test and F-test, considering gender, the results (t-test: p = 0.045, and F-test: p =3.03e-13) reveal that the age distributions of Male and Female subjects are significantly different in terms of Mean and Variance. For ethnicity, results (t-test: p = 2.761e-12, F-test: p = 1.722e-05) show that the age distributions of Asian and Non-Asian subjects are also significantly different. In Fig. 9, we present the age distribution details for each gender and ethnicity. It shows the age distributions vary in different demographic groups, especially between Asian and Non-Asian subgroups. The unbalanced distributions of age for different demographic subgroups may have contributed to the improvement of age estimation results, by injecting more precise prior information in training. While, due to the unavailability of more proper 3D face dataset which covers evenly

Fig. 7 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.

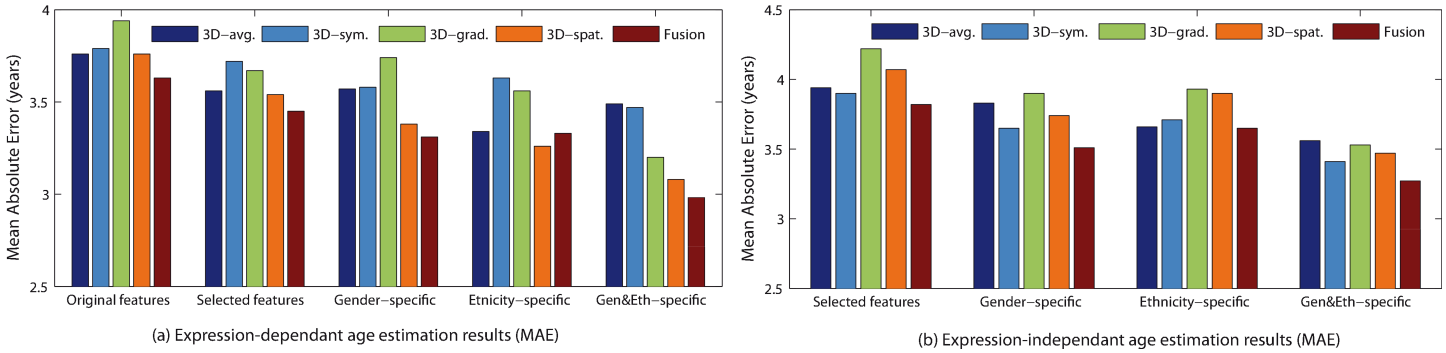
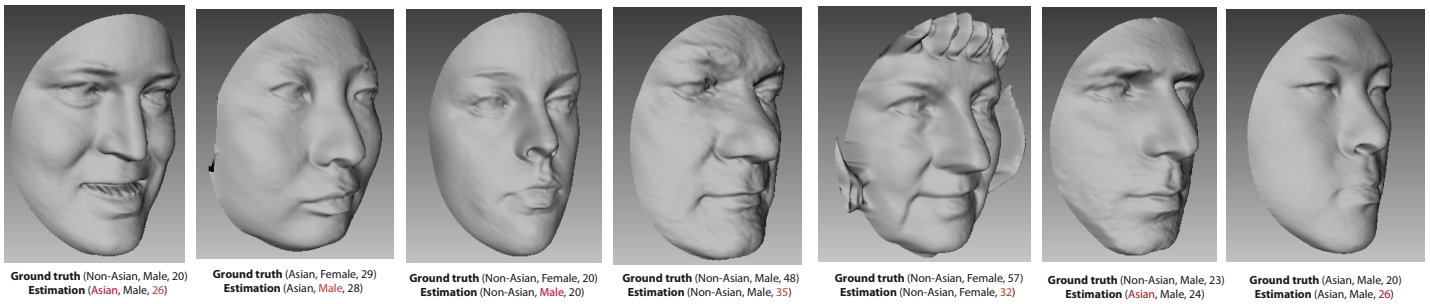


Fig. 8 Examples of 3D faces with the ground truth demographic information and the estimation (in red inaccurate estimated attributes).



the age ranges in different gender and ethnicity subgroup, further study of this effect is not currently feasible. In this case, we note that the *attribute-specific* experimental settings and the statistical tests introduced in the current work are still of value for guiding future research when proper dataset is collected, and the results produced here could also serve as preliminary references of future study. Also, as far as we are concerned, there has been no related work in literature giving explicit statistical analysis on the demographic groups when applying *attribute-specific* settings, even in the related 2D studies [23, 24, 38, 46, 48, 56].

Shown in Fig. 8 are examples of 3D faces with the corresponding ground truth demographic information (ethnicity, gender, age). The estimated attributes are also given for comparison (inaccurate estimations are highlighted in red). Several factors make the estimation hard as the occlusion of the face by the hair or the deformations caused by the expressions, especially when the mouth is open (as nuisance factors). However, the most challenging factor remains the intra-groups variations in term of 3D shape differences.

4.4 Cross-database Validation over FU-3D-Faces

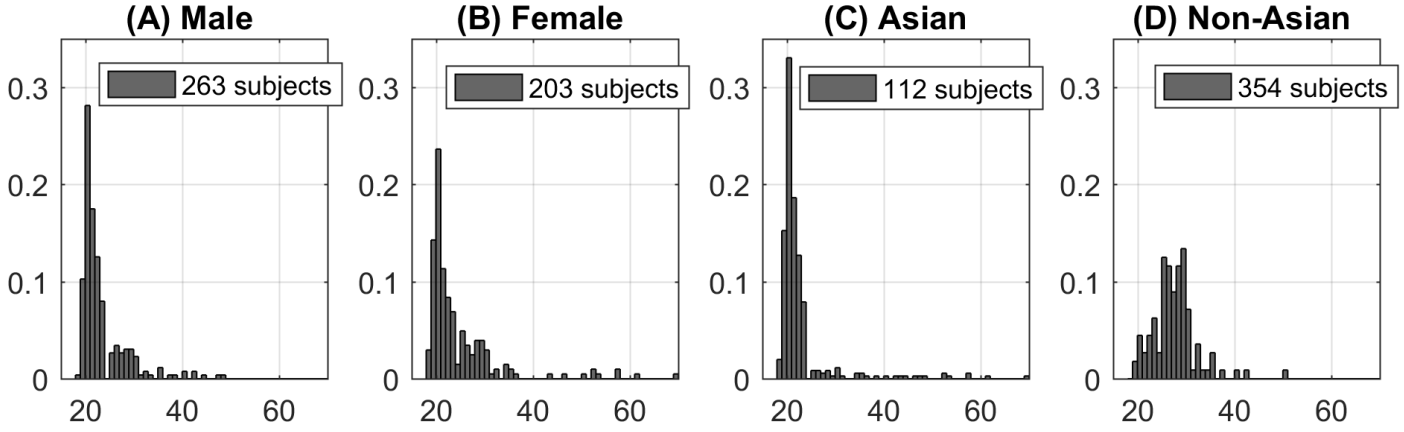
To further evaluate the proposed approach, we have designed another set of gender, ethnicity and age recognition experiments in a cross-dataset settings. Taking the earliest 466 scans of FRGCv2 dataset for training, we perform the recognition on the 53 scans of the Florence 3D face dataset (FU-3D) [4]. The FU-3D dataset is still under collection in the Media Integration and Communication Center at the University of Florence. It contains 53 neutral-frontal 3D scans of 53 Caucasian (Non-Asian) subjects, of which 14 are female and 39 are male. Each subject's age ranges from 22 to 61, with 24 subjects above 30 years old. In these experiments, the scans from the FU-3D are used as sequestered data (never used in the training). The experimental results are shown in Table 2. We point out that even the 3D scanning devices are based on optical triangulation, a laser rangefinder is used to collect faces in FRGCv2 and a multi-cameras (stereo) system is used to capture 3D faces of the FU dataset. The two systems are comparable in terms of resolution. While the face meshes in FRGCv2 are generally smoother, and the FU-3D dataset preserves more local shape changes than FRGCv2.

Table 2 Cross-dataset experiments results (train with the 466 earliest scans of FRGCv2 dataset and test on the FU-3D dataset, with the selected DSF features in each description

	3D-avg.	**3D-sym.**	**3D-grad.**	**3D-spat.**
gender	81.13%	88.67%	83.01%	84. 90%
ethnicity	100%	100%	100%	100%
age (MAE)	5.68	3.60	5.06	6.85

As shown in Table 2, our gender classification method achieved always a classification rate higher than 80% (with the *3D-sym* description, it achieved 88.67%). These results demonstrate that our gender classification approaches have a good generalization performance on the FU-3D-Faces. For

Fig. 9 Age distributions in different gender and ethnicity subgroups of the FRGCv2 dataset. (In each subplot, the horizontal axis denotes the age of subjects, and the vertical axis shows the normalized frequency. The title above each subplot denotes the underlying gender/ethnicity group, and the legend describes the number of subjects in that particular demographic subgroup. Age distribution differences are observed between demographic subgroups, especially for Asian and Non-Asian subjects, shown in subplot (C) and (D))



ethnicity, all the subjects (Caucasians) of the dataset are correctly classified. In age estimation, the best achieved MAE is 3.6 years obtained using the *3D-sym* description. This confirms the strong relationship between the age groups and the facial asymmetry. Overall, the results reported in Table 2 demonstrate the good generalization capability of our recognition methods on other datasets. Because of the small size of the FU-3D dataset, we point out that these results are obtained under age-, gender- and ethnicity-general settings (no joint recognition performed here).

4.5 Comparison to previous work

In the literature, several studies have investigated the problem of demographic traits (gender, ethnicity and age) recognition, with each trait considered separately. The use of the 3D shape is relatively new and the majority of the published papers have reported results on the FRGCv2 dataset [47], the largest dataset which includes the gender, ethnicity and age meta-data. While the tasks of gender and ethnicity classification from 3D facial scans have been explored, this work presents (1) the first study on age estimation using 3D shapes; and (2) the first investigation of joint demographic traits recognition based on 3D shapes. Through the extensive experiments, it brings new valuable conclusions in terms of the correlations between facial soft-biometric traits. In this section, we provide a comparative study of the proposed approaches and previous studies. For age estimation, as shown in Table 3, *Huerta et al.* [31] have performed age estimation experiments on the 2D part of the FRGCv2 dataset. Following a 5-fold cross-validation setting, they achieved 4.17 years MAE. In this work, we achieve a much lower MAE of 3.24 years. We point out the following differences between the 2D and 3D parts of FRGCv2 dataset – (1) 2D part contains 100 subjects more than the 3D part; and (2) Number of images in the 3D part (4,007) is 10 times lower than in the 2D part (44,278). Considering these differences, this comparison demonstrates the interest of using the proposed 3D morphology-driven features in age estimation.

Several recent studies have proposed 3D-based gender and ethnicity classification, as shown in Table 4 and Table 5. The results are reported mainly under the 10-Fold Cross-Validation ***(10-CV)*** setting, for which the subjects are first partitioned equally into 10 groups, then each group is used for testing once with the rest 9 groups used in training. To compare directly with the state-of-the-art, in Table 4 and Table 5, we generate and present the Ethnicity&Age-specific gender classification results, and the Gender&Age-specific ethnicity classification results, under the ***(10-CV)*** protocol. As shown in Table 4, for 3D gender classification, the works closely related to ours are presented in [54], [5], [21], [59], which are also tested on the FRGCv2 dataset. Our gender classification rate on the 466 earliest scans of FRGCv2 (95.28%) outperforms significantly the results of *Ballihi et al.* in [5] (86.05%). Our result is 1.77% lower than *Gilani et al.* in [21] (97.05%). With all the 4,007 scans of FRGCv2, we achieve 93.61% gender classification rate. This results is comparable to the results of *Toderici et al.* in [54] (93.5%), *Wang et al.* in [59] (93.7%), and is 2.51% lower than the result of *Gilani et al.* in [21] (96.12%). Overall, the proposed approach achieves competitive results with existing works. The approach of *Gilani et al.* in [21] based on 3D landmarks on the face outperforms remaining approaches (including ours), however, their model depends on the accuracy of 3D the landmarks localization algorithm. For ethnicity classification, as shown in Table 5, the most comparable works are proposed by Zhong et al. in [63], and *Toderici et al.* in [54]. With the 466 earliest scans of FRGCv2, we achieve 95.92% ethnicity classification rate following the 10-fold cross-validation protocol. Using all the 4,007 scans of FRGCv2, we achieve 96.60% ethnicity classification rate. This result is higher than the result of *Zhong et al.* in [63] (82.38%). Compared to *Toderici et al.* in [54], our ethnicity classification rate is 2.4% lower. However, their result was reported on only the 3676 Asian and White scans of FRGCv2. The scans of the remaining subjects were 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. Thus, compared to the

Table 3 Comparison of age estimation results with [31]. (Experiments conducted on 3D and 2D parts of FRGCv2 dataset, respectively.)

Authors	Imagery	Method	Regressor	Database	# of subjects/images	MAE (years)
Huerta et al. [31]	*2D*	*HOG+GRAD+LBP+SURF*	*CCA*	*FRGC (2D part)*	*568/44,278*	*4.17*
This work	*3D*	*SYM+AVE+GRAD+SPAT*	*Random Forest*	*FRGC (3D part)*	*466/4,007*	***3.24***

Table 4 Comparison of the proposed gender classification approach with state-of-the-art.

Author	Dataset	Auto	Features	Classifiers	Setting	Results	Imagery
Ballihi et al. [5]	*466 scans of FRGCv2*	*Yes*	*Facial curves*	*Adaboost*	*10-CV*	86.05%	*3D*
Toderici et al. [54]	*3,675 scans of FRGCv2*	*Yes*	*Wavelets*	*Polynomial-SVM*	*10-CV*	93.5%	*3D*
Huang et al. [30]	*3,676 scans of FRGCv2*	*Yes*	*LCP*	*Adaboost*	*10-CV*	95.5%	*2D+3D*
Wang et al. [59]	*4,007 scans of FRGCv2*	*No*	*3D coordinates*	*RBF-SVM*	*5-CV*	*93.7%*	*2D+3D*
Gilani et al. [21]	*466 scans of FRGCv2*	*Yes*	*landmark distances*	*LDA classifier*	*10-CV*	*97.05%*	*3D*
	4,007 scans of FRGCv2	*Yes*	*landmark distances*	*LDA classifier*	*10-CV*	*96.12%*	*3D*
This work	*466 scans of FRGCv2*	*Yes*	*3D features*	*Random Forest*	*10-CV*	***95.28%***	*3D*
	4,007 scans of FRGCv2	*Yes*	*3D features*	*Random Forest*	*10-CV*	***93.61%***	*3D*

Table 5 Comparison of the proposed ethnicity classification approach with state-of-the-ar.

Author	Dataset	Auto	Features	Classifiers	Setting	Results	Imagery
Zhong et al. [63]	*4,007 scans of FRGCv2*	*No*	*LVC features*	*Class Probability*	*No-CV*	82.38%	*3D*
Lu et al. [44]	*Subset of UND and MSU*	*No*	*Grid elements*	*SVM*	*10-CV*	*98%*	*3D + 2D*
Toderici et al. [54]	*3,676 scans of FRGCv2*	*Yes*	*Wavelets*	*Polynomial-SVM*	*10-CV*	99.5%	*3D*
Huang et al. [30]	*3,676 scans of FRGCv2*	*Yes*	*LCP*	*Adaboost*	*10-CV*	99.6%	*2D+3D*
This work	*466 scans of FRGCv2*	*Yes*	*3D features*	*Random Forest*	*10-CV*	***95.92%***	*3D*
	4,007 scans of FRGCv2	*Yes*	*DSF features*	*Random Forest*	*10-CV*	***96.60%***	*3D*

work of *Toderici et al.*, we have encountered significantly more complicated ethnicity challenges.

Based on the comparative study presented above, we have presented several contributions – (1) We have proposed novel morphology-driven features extracted from 3D faces, they are designed based on a number of studies in the biological and clinical studies. We underline the novelty of these 3D features compared to the widely used conventional features used in 2D face analysis, such as the BIF (Bio-inspired Features), or the texture features (Gabor, LBP, Haar, etc.); (2) The proposed framework is quite generic to be used in individual (or joint) age, gender and ethnicity recognition. We have demonstrated that considering the correlations between the demographic traits can improve the recognition results than taking them individually, thus reveal the interest of estimating them jointly; (3) We have demonstrated through the several presented experiments the effectiveness of the proposed approach and its competitive results with the state-of-the-art. In addition, processing 3D facial meshes is often time consuming compared to processing intensity or color images. However, through the approximation by a collection of 3D radial curves and the use of an efficient shape analysis framework [51], developed in C++, the proposed approach shows time efficiency in 3D curves extraction and feature extraction as presented in Table 6. All our experiments are performed on an Intel Core 2 Duo 2.53 GHz processor, with 4 GB of memory. This demonstrates the effectiveness of the processing pipeline for 3D feature extraction for soft-biometric traits recognition.

Table 6 Processing time for face pre-processing and features extraction.

Step	Details	time (s)
Pre-processing	holes filling, cropping, smoothing	< 1
Curves extraction	200, 100 vertex/curve	0.8
3D features	shape analysis	4×0.05
Overall	4 × 20K 3D shape features	**2**

5 Conclusions

This paper presents a set of new morphology-driven features extracted from the 3D shape of the face and investigates the joint demographic (gender, ethnicity and age) estimation based on them. It provides also a comprehensive study on their relevance and highlights the most informative areas of the 3D face for age, gender and ethnicity. The proposed 3D features are used individually (and fused) to perform in the same time gender, ethnicity and age estimation which makes them applicable for the three problems as well as the joint estimation. Results reported on FRGCv2 (following LOPO and 10-fold cross validation settings) are competitive compared to the state-of-the-art with the advantage that the proposed approach is fully automatic and presents interesting execution time. The cross-database validation experiment conducted on the FU-3D-Faces dataset confirms the obtained good results. The problem of age estimation from the 3D shape of the face is investigated for the first time in the literature. However, the datasets used in the study present some limitations – (1) compared to some 2D face datasets recently used to evaluate the age estimation algorithms such as MORPH II and PCSO [26], the 3D datasets are limited in number of subjects and images; (2) the age distributions are biased as the majority of participants were Caucasian students. We note in particular that, at the current stage, the unbalanced distribution of subjects' age in different gender of ethnicity, remains an unmeasured confounding factor for demonstrating the usefulness of *attribute-specific* age estima-

tion. Thus, one important item to promote 3D-based demographic information estimation (including the age) will be to collect new 3D face datasets to handle these limitations. Also, when proper 3D face dataset is available, quantitative studying these soft-biometrics' correlations would also be of research interest, which could promote the understanding of their co-existence and co-exhibition in 3D faces.

Acknowledgements This work was supported by the ANR project 2010 INTB 0301 01, the CMCU project number 34882WK, the Ph.D. scholarship from the Chinese Scholarship Council (CSC) to Baiqiang Xia and partially supported by the FUI project MAGNUM 2. The authors would like to thank the colleagues at the Media Integration and Communication Center (MICC) of the University Florence for providing the FU-3D-Faces dataset.

References

1. Alice J, O., Thomas, V., Harald, V., Elizabeth M, S.: Three-dimensional caricatures of human heads: Distinctiveness and the perception of facial age. PERCEPTION **26**, 719–732 (1997)
2. Alphonse, J., Cox, J., Clarke, J., Schluter, P., McLennan, A.: The effect of ethnicity on 2D and 3D frontomaxillary facial angle measurement in the first trimester. Obstetrics and Gynecology International (2013)
3. Ashok, S., Vanitha, S., David, M.: Analysis of sexual dimorphism in human face. Journal of Visual Communication and Image Representation **18**, 453–463 (2007)
4. Bagdanov, A.D., Del Bimbo, A., Masi, I.: The florence 2D/3D hybrid face dataset. In: Proceedings of the 2011 Joint ACM Workshop on Human Gesture and Behavior Understanding, J-HGBU '11, pp. 79–80 (2011)
5. Ballihi, L., Ben Amor, B., Daoudi, M., Srivastava, A., Aboutajdine, D.: Boosting 3-D geometric features for efficient face recognition and gender classification. IEEE Trans. on Information Forensics and Security **7**(6), 1766–1779 (2012)
6. Baudouin, J.Y., Gallay, M.: Is face distinctiveness gender based? Journal of Experimental Psychology: Human Perception and Performance **32**, 789 (2006)
7. Ben Amor, B., Drira, H., Berretti, S., Daoudi, M., Srivastava, A.: 4D facial expression recognition by learning geometric deformations. IEEE Trans. Cybernetics **44**(12), 2443–2457 (2014)
8. Breiman, L.: Random forests. Machine Learning **45**(1), 5–32 (2001)
9. Bruce, V., Burton, A.M., Hanna, E., Healey, P., Mason, O., Coombes, A.: Sex discrimination: How do we tell the difference between male and female faces? Perception **22(2)**, 131–152 (1993)
10. Clinton S, M., Benjamin Z, P., Johnny T, C., Stephen R, S., Helena OB, T.: The relationship between age and facial asymmetry (2011). URL http://meeting.nesps.org/2011/80.cgi
11. Dantcheva, A., Elia, P., Ross, A.: What else does your biometric data reveal? a survey on soft biometrics. IEEE Transactions on Information Forensics and Security **11**(3), 441–467 (2016). DOI 10.1109/TIFS.2015.2480381
12. Deffenbacher, K.A., Vetter, T., Johanson, J., O'Toole, A.J.: Facial aging, attractiveness, and distinctiveness. PERCEPTION **27**, 1233–1244 (1998)
13. Drira, H., Ben Amor, B., Srivastava, A., Daoudi, M., Slama, R.: 3D face recognition under expressions, occlusions, and pose variations. IEEE Trans. Pattern Anal. Mach. Intell. **35**(9), 2270–2283 (2013)
14. Enlow, D.H.: A morphogenetic analysis of facial growth. American journal of orthodontics **52**, 283–299 (1966)
15. Farinella, G., Dugelay, J.: Demographic classification: Do gender and ethnicity affect each other? In: Informatics, Electronics Vision (ICIEV), 2012 International Conference on, pp. 383–390 (2012)
16. Farkas, L.G., Katic, M.J., Forrest, C.R.: International anthropometric study of facial morphology in various ethnic groups/races. Journal of Craniofacial Surgery **16**(4), 615–646 (2005)
17. Fu, S., He, H., Hou, Z.G.: Learning race from face: A survey. IEEE Transactions on Pattern Analysis and Machine Intelligence **36**(12), 2483–2509 (2014). DOI 10.1109/TPAMI.2014.2321570
18. Fu, Y., Guo, G., Huang, T.S.: Age synthesis and estimation via faces: A survey. Pattern Analysis and Machine Intelligence, IEEE Transactions on **32**(11), 1955–1976 (2010)
19. Gao, W., Ai, H.: Face gender classification on consumer images in a multiethnic environment. In: Advances in Biometrics, *Lecture Notes in Computer Science*, vol. 5558, pp. 169–178 (2009)
20. Geng, X., Zhou, Z., Smith-Miles, K.: Automatic age estimation based on facial aging patterns. IEEE Trans. Pattern Anal. Mach. Intell. **29**(12), 2234–2240 (2007)
21. Gilani, S.Z., Shafait, F., Mian, A.S.: Biologically significant facial landmarks: How significant are they for gender classification? In: 2013 International Conference on Digital Image Computing: Techniques and Applications, pp. 1–8 (2013)
22. Guo, G.: Human age estimation and sex classification pp. 101–131 (2012)
23. Guo, G., Fu, Y., Dyer, C.R., Huang, T.S.: Image-based human age estimation by manifold learning and locally adjusted robust regression. IEEE Trans. on Image Processing **17**(7), 1178–1188 (2008)
24. Guodong, G., Guowang, M.: Human age estimation: What is the influence across race and gender? In: Computer Vision and Pattern Recognition Workshops (CVPRW), pp. 71–78 (2010)
25. Hall, M.A.: Correlation-based feature subset selection for machine learning. Ph.D. thesis, University of Waikato (1999)
26. Han, H., Otto, C., Liu, X., Jain, A.: Demographic estimation from face images: Human vs. machine performance. IEEE Trans. Pattern Anal. Mach. Intell. **PP**(99), 1–1 (2014)
27. Han, X., Ugail, H., Palmer, I.: Gender classification based on 3D face geometry features using svm. In: 2013 International Conference on Cyberworlds, pp. 114–118 (2009)
28. Hayashi, J.i., Yasumoto, M., Ito, H., Koshimizu, H.: Age and gender estimation based on wrinkle texture and color of facial images. In: 16th International Conference on Pattern Recognition, vol. 1, pp. 405–408. IEEE (2002)
29. Hu, Y., Fu, Y., Tariq, U., Huang, T.S.: Subjective experiments on gender and ethnicity recognition from different face representations. In: Advances in Multimedia Modeling, 16th International Multimedia Modeling Conference, pp. 66–75 (2010)
30. Huang, D., Ding, H., Wang, C., Wang, Y., Zhang, G., Chen, L.: Local circular patterns for multi-modal facial gender and ethnicity classification. Image and Vision Computing **32**(12), 1181 – 1193 (2014)
31. Huerta, I., Fernandez, C., Prati, A.: Facial age estimation through the fusion of texture and local appearance descriptor. In: IEEE European Conference on Computer Vision (ECCV 2014) - Int. Workshop on Soft Biometrics (2014)
32. Hunter, W.S., Garn, S.M.: Disproportionate sexual dimorphism in the human face. American Journal of Physical Anthropology **36**, 133–138 (1972)
33. Huynh, T., Min, R., Dugelay, J.: An efficient LBP-based descriptor for facial depth images applied to gender recognition using RGB-D face data. In: ACCV 2012 International Workshops on Computer Vision, pp. 133–145 (2012)
34. Jain, A.K., Dass, S.C., Nandakumar, K.: Soft biometric traits for personal recognition systems. In: First International Conference on Biometric Authentication, ICBA 2004, pp. 731–738 (2004)
35. Jennifer, A., Jennifer, C., Jill, C., Philip, S., Andrew, M.: The effect of ethnicity on 2D and 3D frontomaxillary facial angle measurement in the first trimester. In: Obstetrics and Gynecology International (2013)
36. Joseph W, Y.: Head and face anthropometry of adult US civilians (1993)
37. Kohavi, R.: Wrappers for performance enhancement and oblivious decision graphs. Ph.D. thesis, Stanford University (1995)
38. Kumar, N., Berg, A.C., Belhumeur, P.N., Nayar, S.K.: Describable visual attributes for face verification and image search. IEEE Trans. Pattern Anal. Mach. Intell. **33**(10), 1962–1977 (2011)
39. Lanitis, A.: On the significance of different facial parts for automatic age estimation. In: Digital Signal Processing, 14th International Conference on, vol. 2, pp. 1027–1030. IEEE (2002)
40. Leslie G, F., Marko J, K., Christopher R, F.: International anthropometric study of facial morphology in various ethnic groups/races. Journal of Craniofacial Surgery **16**, 615–646 (2005)

41. Little, A.C., Jones, B.C., DeBruine, L.M., Feinberg, D.R.: Symmetry and sexual dimorphism in human faces: interrelated preferences suggest both signal quality. Behavioral Ecology **19**, 902–908 (2008)
42. Little, A.C., Jones, B.C., Waitt, C., Tiddeman, B.P., Feinberg, D.R., Perrett, D.I., Apicella, C.L., Marlowe, F.W.: Symmetry is related to sexual dimorphism in faces: Data across culture and species **3** (2008)
43. Liu, Y., Palmer, J.: A quantified study of facial asymmetry in 3D faces. In: IEEE International Workshop on Analysis and Modeling of Faces and Gestures, AMFG '03, pp. 222– (2003)
44. Lu, X., Chen, H., Jain, A.K.: Multimodal facial gender and ethnicity identification. In: Advances in Biometrics, International Conference, ICB 2006, pp. 554–561 (2006)
45. Ng, C.B., Tay, Y.H., Goi, B.M.: Vision-based human gender recognition: A survey. arXiv preprint arXiv:1204.1611 (2012)
46. N.S., L., J., B., S., M.: Age estimation using gender information. In: K. Venugopal, L. Patnaik (eds.) Computer Networks and Intelligent Computing, *Communications in Computer and Information Science*, vol. 157, pp. 211–216 (2011)
47. Phillips, P.J., Flynn, P.J., Scruggs, W.T., Bowyer, K.W., Chang, J., Hoffman, K., Marques, J., Min, J., Worek, W.J.: Overview of the face recognition grand challenge. In: 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2005), pp. 947–954 (2005)
48. Ramanathan, N., Chellappa, R., Biswas, S.: Computational methods for modeling facial aging: A survey. Journal of Visual Languages & Computing **20**(3), 131 – 144 (2009)
49. Rhodes, M.G.: Age estimation of faces: A review. Applied Cognitive Psychology **23**(1), 1–12 (2009)
50. Shirakabe, Y., Suzuki, Y., Lam, S.: A new paradigm for the aging asian face. Aesthetic Plastic Surgery **27**(5), 397–402 (2003)
51. Srivastava, A., Klassen, E., Joshi, S.H., Jermyn, I.H.: Shape analysis of elastic curves in euclidean spaces. IEEE Trans. Pattern Anal. Mach. Intell. **33**(7), 1415–1428 (2011)
52. Steven, W., Randy, T.: Facial masculinity and fluctuating asymmetry. Evolution and Human Behavior **24**, 231–241 (2003)
53. Suo, J., Zhu, S.C., Shan, S., Chen, X.: A compositional and dynamic model for face aging (2010)
54. Toderici, G., O'Malley, S.M., Passalis, G., Theoharis, T., Kakadiaris, I.A.: Ethnicity- and gender-based subject retrieval using 3-D face recognition techniques. International Journal of Computer Vision **89**(2-3), 382–391 (2010)
55. Tokola, R., Mikkilineni, A., Boehnen, C.: 3d face analysis for demographic biometrics. In: Biometrics (ICB), 2015 International Conference on, pp. 201–207. IEEE (2015)
56. Ueki, K., Sugiyama, M., Ihara, Y.: Perceived age estimation under lighting condition change by covariate shift adaptation. In: 20th International Conference on Pattern Recognition, ICPR 2010, pp. 3400–3403 (2010)
57. Vicki, B., A Mike, B., Elias, H., Pat, H., Oli, M., Anne, C., Rick, F., Alf, L.: Sex discrimination: How do we tell the difference between male and female faces? Perception **22**, 131–152 (1993)
58. Vignali, G., Hill, H., Vatikiotis-Bateson, E.: Linking the structure and perception of 3D faces: Gender, ethnicity, and expressive posture. In: AVSP 2003 - International Conference on Audio-Visual Speech Processing, St. Jorioz, France, September 4-7, 2003, pp. 193–198 (2003)
59. Wang, X., Kambhamettu, C.: Gender classification of depth images based on shape and texture analysis. In: IEEE Global Conference on Signal and Information Processing, GlobalSIP 2013, Austin, TX, USA, December 3-5, 2013, pp. 1077–1080 (2013)
60. Wu, J., Smith, W.A.P., Hancock, E.R.: Facial gender classification using shape-from-shading. Image Vision Comput. **28**(6), 1039–1048 (2010)
61. Xia, B., Ben Amor, B., Daoudi, M., Drira, H.: Can 3D shape of the face reveal your age? In: VISAPP 2014 - Proceedings of the 9th International Conference on Computer Vision Theory and Applications, pp. 5–13 (2014)
62. Xia, B., Ben Amor, B., Huang, D., Daoudi, M., Wang, Y., Drira, H.: Enhancing gender classification by combining 3D and 2D face modalities. In: 21st European Signal Processing Conference, EUSIPCO 2013, pp. 1–5 (2013)
63. Zhong, C., Sun, Z., Tan, T.: Fuzzy 3D face ethnicity categorization. In: Advances in Biometrics, Third International Conference, ICB 2009, Alghero, Italy, June 2-5, 2009. Proceedings, pp. 386–393 (2009)
64. Zhuang, Z., Bradtmiller, B.: Head and face anthropometric survey of us respirator users **2**, 567–576 (2005)
65. Zhuang, Z., Landsittel, D., Benson, S., Roberge, R., Shaffer, R.: Facial anthropometric differences among gender, ethnicity, and age groups. Annals of Occupational Hygiene (2010)