Ethnicity Classification by the 3D Discrete Landmarks Model measure in Kendall shape space

Article *in* Pattern Recognition Letters · October 2019

DOI: 10.1016/j.patrec.2019.10.035

Pattern Recognition Letters 129 (2020) 26–32

Contents lists available at ScienceDirect

Pattern Recognition Letters

journal homepage: www.elsevier.com/locate/patrec

Ethnicity classification by the 3D Discrete Landmarks Model measure in Kendall shape space☆

Chenlei Lv, Zhongke Wu, Xingce Wang*, Zhang Dan, Mingquan Zhou

School of Artificial Intelligence, Beijing Normal University, 100875, China

A R T I C L E I N F O

Article history:
Received 24 March 2018
Revised 22 October 2019
Accepted 30 October 2019
Available online 31 October 2019

Keywords:
Ethnicity classification
Facial Landmark Model
Kendall shape space,

A B S T R A C T

Knowledge of human ethnicity constitutes important biometric information. An automated ethnicity classification is a good first step in facial analysis. However, most ethnicity classification methods require a complex feature extraction and model training process. We propose a novel ethnicity classification method based on the analysis of facial landmarks in Kendall shape space. Facial features with different relative positions have a close relationship with ethnicity. Facial landmarks can represent positions of facial features. We build a Discrete Landmarks Model (DLM) based on facial landmarks and construct an ethnicity classification model based on the DLM analysis. The clear advantages of our method are that it is fully automated; requires no complex data preprocessing, feature extraction or a complex training process; results in a fast and accurate classification process. We estimate the effectiveness of our method experimentally, using public databases such as Texas3D, FRGC2.0, BU-3DFE and BU-4DFE. On average, our method can achieve a 95% ethnicity classification rate with each classification attempt in 2.0 s.

1. Introduction

Ethnicity attributes of humans provide important information for anthropology research and automated analysis of facial data. Humans normally rely on vision to recognize ethnicity. However, ethnicity classification in an automated framework is challenging when using 2D images. This is caused by the 2D imagery-based ethnicity classification process being affected by numerous factors such as illumination, use of make-up and variations in head poses. In comparison, an ethnicity classification process based on 3D facial data can achieve more accurate results. It is robust to texture variations. Our ethnicity classification framework uses 3D facial data. Certain challenges arise when building an automated ethnicity classification framework using such data: the data quality is limited by various kinds of 3D scanning equipment; a raw 3D scanning data reconstruction process generally requires manual intervention; and feature extraction based on mesh analysis is complex.

To overcome such challenges, we propose a novel automated ethnicity classification method based on 3D facial landmarks data. The advantages of our method are clear. First, facial landmark detection has been studied for many years, resulting in numerous mature algorithms [26,32]. Second, with the development of binocular cameras, 3D facial landmark data can be easily computed by combining 2D facial landmarks. Finally, we solely use positions of 3D landmarks to build the classification model, without reliance on 2D imagery or 3D surface information. Hence, the feature extraction process in our method is extremely concise. The relative positions of facial features can be measured by 3D landmarks. In summary, ethnicity classification can be performed via 3D landmarks modeling and measurement.

Our method includes three steps: detection of 3D landmarks, construction and measurement of the Discrete Landmarks Model (DLM), and Discrete Landmarks Model-based clustering and classification. The framework is shown in Fig. 1. In conclusion, the contributions of our method are as follows:

1. We propose a representation of 3D landmarks called the Discrete Landmarks Model (DLM). The DLM can represent a distribution of facial features that has a close relationship with human perception of facial data.

2. We propose a measure for DLMs in the Kendall shape space. The DLMs are embedded in the Kendall shape space. The difference between different DLMs can be represented by geodesic distances in the space. The measure result is robust to facial expressions, which is based on the DLM reconstruction.

3. We propose an ethnicity classification method based on the DLMs' measures. The classification process does not require complex feature extraction or data training. The classification process is fast.

☆ Editor: Prof. S. Sarkar.

* Corresponding author
E-mail address: wangxingce@bnu.edu.cn (X. Wang).

https://doi.org/10.1016/j.patrec.2019.10.035
0167-8655/

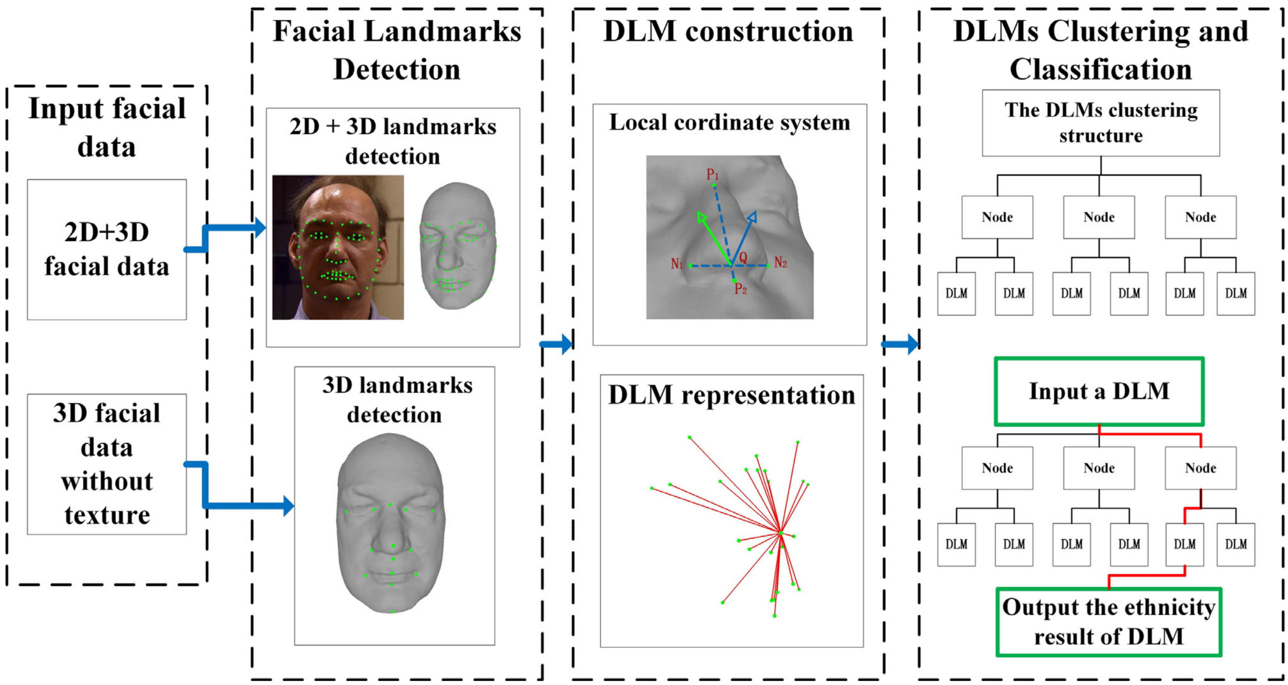

Fig. 1. The pipeline for constructing the ethnicity classifier.

The remaining parts of the paper are organized as follows. In Section 2, we review the related work on facial classification and facial landmark models. In Section 3, we illustrate 3D landmarks detection in different types of facial data and construct the DLM. In Section 4, we propose a measure for a DLM in the Kendall shape space and cluster the DLMs to create an ethnicity classification framework. In Section 5, we apply our method to publicly available facial databases, Texas3D, FRGC2.0, BU-3DFE and BU-4DFE.

2. Related works

Facial classification has been studied for many years. The classification tasks include identifying the subject's ethnicity, gender, age and facial expressions. The methods of gender and ethnicity classification research can be divided into two parts: texture-based and geometric feature-based.

A classification framework using a texture-based approach relies on texture information from a 2D facial image. The classification framework extracts features from facial images to build a classifier. The relevant image features include binary features [1], biologically-inspired features [2], demographic features [3], the Gabor filter process features [4,5] and local binary patterns (LBP) [6–8]. The classical learning frameworks that are also required to analyze features include Support Vector Machine (SVM) [9,10], Linear Discriminant Analysis (LDA) [11], Principal Component Analysis (PCA) [12,13], random forest [8,13], AdaBoost [14] and the convolutional neural network (CNN) [15]. Using the texture information in the classification task has several advantages: the source facial images can be obtained by readily available hardware; and it is easy to build an analysis model using a large quantity of data and set up online processing. However, the analysis of image features is affected by certain factors, such as hair occlusion, extremes of lighting, motion blur and variations in head poses.

Geometry-based approaches extract geometric features from facial data. In most cases, facial data are obtained from 3D imagery. As a result, 3D facial data retain the full spatial information. Geometric features include mesh features [16], profiles and curvature [17], radial and isolevel curves [18,19,21], 3D landmarks-based geodesic distance [22], symmetry properties [20] and correspondence vectors [23]. Using geometric features in classification tasks was robust to texture variations. However, 3D scanning equipment is uncommon in everyday life. 3D data are difficult to obtain for ordinary users. A mesh reconstruction was necessary to extract geometric features of an accurate surface. The mesh reconstruction algorithms' use of raw data also limits the widespread use of geometry-based approaches.

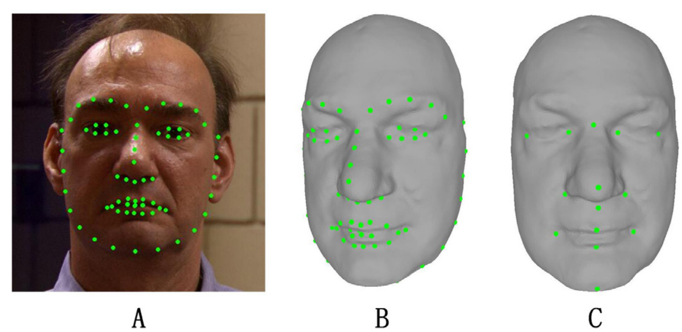

Fig. 2. Examples of various methods for detecting facial landmarks (A: 2D facial landmarks detection; B: a 3D representation of 2D landmarks on a 3D surface; C: detection of 3D facial landmarks directly from a 3D facial surface).

Generally, methods based on 3D geometric feature analysis can achieve superior classification results. We construct an ethnicity classification framework by utilizing 3D facial landmarks data. Detection of facial landmarks has been studied for many years, using 2D imagery [32] and 3D facial surfaces [26,33–35]. Typically, 3D scanning devices output 3D facial data together with the corresponding 2D facial image. The position of each pixel in the 2D image corresponds to a related point in the 3D facial data. In our framework, we obtain positions of 3D landmarks from the mapping of landmarks' positions. In Fig. 2, we show examples of several landmark detection methods.

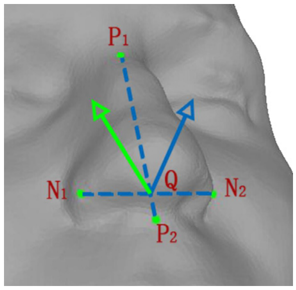

Fig. 3. An illustration of a local coordinate system created from several landmarks. P_1, P_2, N_1, N_2 are landmarks around the nose area. Q is the vector product of P_1P_2 and Q, which is computed by $P_1P_2{}^*N_1N_2$.

3. Discrete Landmarks Model (DLM)

Our ethnicity classification framework is based on 3D facial landmarks data. Based on facial landmarks, the landmarks model can be constructed, such as the Active Shape Model (ASM) [24], the Active Appearance Model (AAM) [25] and the facial landmark model (FLM) [26]. We construct a sample model, called the Discrete Landmarks Model (DLM), of landmarks for facial data analysis. The DLM is used to model landmarks in a way that is invariant to rigid transformations in R^3. The construction of the DLM includes three parts: (1) removal of translation, (2) removal of scale, and (3) rotation alignment. Eqs. (1)–(3) show the construction of the DLM.

$$F = \{x_1, \dots x_k\}, x_B \in F; L = \{v_1, \dots v_k\}, v_i = x_i - x_B \tag{1}$$

In Eq. (1), F is the set of facial landmarks, while x_B is the benchmark of F. The benchmark is used to normalize locations of landmarks on different faces. We select the nose tip landmark as the benchmark. L is the set of vectors of all landmarks relative to the benchmark. As a result, we obtain the preliminary landmarks model.

$$s(L) = \left(\sum_{i=1}^{k} \|v_i\|\right)$$
$$L_s = \left\{v'_1, \dots v'_k\right\}, v'_i = \frac{v_i}{s(L)}, \sum_{i=1}^{k} \left\|v'_i\right\| = 1 \tag{2}$$

In Eq. (2), s is the scale of L. Different scales of L influence the analysis of relative positions of facial features. The landmarks model should have a standardized scale. To remove the scaling factor, we divide each vector in L by s. The result is the new landmarks model L_s. The scale of L_s is unitized, while the relative positions of vectors in L remain unchanged.

$$L_{sr} = \{T(v'_1), \dots, T(v'_k)\} \tag{3}$$

In Eq. (3), we obtain L_{sr}, the final representation of the DLM. The vectors in L_s are still affected by the rotation factor. We introduce a transform function T to remove rotation from the landmarks model. Function T is a coordinate transformation based on a local coordinate system. We select the landmarks around the nose to build the coordinate system which is robust to facial expressions. Fig. 3 shows an example of a local coordinate system.

4. Ethnicity classification framework

The DLM is used to represent relative positions of facial features in 3D facial data. Our ethnicity classification framework is based on the DLM. The framework includes three steps: DLM reconstruction, DLM measurement and the classifier construction. The influence of facial expression should be removed from DLM. The DLM reconstruction is used to reduce such influence. The measurement of DLM is computed in the Kendall shape space [30,31]. Based on the DLM measurement, the Ethnicity classifier is constructed.

4.1. DLM reconstruction

The influence of expressions should be considered for DLM measurement. We propose DLM reconstruction to reduce the expression influence. The DLM reconstruction can be regarded as linear optimization process. Based on a template facial database, we achieve a reconstruct DLM to match the original DLM, which have removed the facial expression. Following two optimization directions (identity and expression), the linear optimization process synthesis the reconstruct DLM. In Eq. (4), we show the linear optimization process.

$$\begin{aligned} &F_t = \{L_{sr}(f_{11}), \dots, L_{sr}(f_{nn})\} \\ &L_{sr}(a)_i = linear\{L_{sr}(a), [L_{sr}(f_{i1}), \dots, L_{sr}(f_{in})]\} \\ &L_{sr}(a)_{ij} = \min\{L_{sr}(a)_i, i \in [1, n]\} \end{aligned} \tag{4}$$

F_t represents the DLM set from template facial database. f_{ij} is the facial data in the database, i represents the index of expression, j represents the index of identity. $L_{sr}(a)_i$ is the linear optimization result of $L_{sr}(a)$ from the DLM subset, which have same expression index. $L_{sr}(a)_{ij}$ is the final linear optimization result. The $L_{sr}(a)_{ij}$ can be regarded as the DLM reconstruct result. To remove the influence of facial expressions, we adjust the weight of optimization result to neutral expression index. In Eq. (5), we show the expression remove process. w_i is the weight of optimization result. The expression index u represents the neutral expression index. We use the $L_{sr}(f_{un})$ to replace $L_{sr}(f_{in})$, the influence of facial expression can be removed. $L_{sr}(a)_{new}$ is the final DLM after reconstruction.

$$\begin{aligned} &L_{sr}(a)_{ij} = w_1 L_{sr}(f_{i1}) + \dots + w_n L_{sr}(f_{in}) \\ &L_{sr}(a)_{new} = w_1 L_{sr}(f_{u1}) + \dots + w_n L_{sr}(f_{un}) \end{aligned} \tag{5}$$

4.2. DLM measurement

Based on the reconstruct DLM, we propose the DLM measurement. In Kendall's theory, the shape of a set of discrete points can be measured by the geodesic distance in the manifold space called the Kendall shape space. We use the corresponding equation to obtain the DLM measurement.

$$d(L_{sr}(a), L_{sr}(b))) = \arccos(L_{sr}(a)_{new} \cdot L_{sr}(b)_{new}) \tag{6}$$

In Eq. (6), a and b represent the facial landmark sets derived from two different facial samples. $L_{sr}(a)$ and $L_{sr}(b)$ are DLM forms for a and b, respectively. In the DLM construction, the influence of symmetry transformation group (translation, scaling and rotation) have been removed. The DLM set belongs to a submanifold of the Kendall shape space. It can be thought of as a multidimensional sphere. The geodesic distance between DLMs is the arc length in spherical coordinates. It is similar to an angle between vectors in R_3. According to Kendall's theory, the transform of the geodesic path between two DLMs in spherical coordinates can be represented by Eq. (7).

$$\begin{aligned} \alpha(k) &= \frac{1}{\sin(\theta)}(\sin(\theta(1-k))L_{sr}(a) + \sin(\theta k)L_{sr}(b))\theta \\ &= d(L_{sr}(a), L_{sr}(b)), k \in [0, 1] \end{aligned} \tag{7}$$

The $a(k)$ is the DLM along the geodesic path between $L_{sr}(a)$ and $L_{sr}(b)$, while θ is the included angle between $L_{sr}(a)$ and $L_{sr}(b)$. The result also represents the geodesic distance between $L_{sr}(a)$ and $L_{sr}(b)$. Parameter k represents a position along the geodesic path.

4.3. The ethnicity classifier construction

Our ethnicity classifier is based on the measurement of DLMs. Based on the anthropology, the similarity of facial features is closely related to ethnic background. The DLMs from the same subclass have high probability of sharing the same ethnic background. Based on the property,we construct the ethnicity classifier. First, we cluster DLMs (by Frey and Dueck [27]) from a facial database to obtain a two-level structure. It is a tree structure used to organize DLMs. The DLMs in the structure are divided into several subclasses. The clustering process is based on the DLM measurement. Second, we design the ethnicity classification function according to the DLM structure. The function is constructed by the analysis of DLMs in the respective subclass.

$$C_{min}(L_{sr}(t)) = \{C_i | C_i \in S, min\{d(L_{sr}(c_i), L_{sr}(t))\}\}. \tag{8}$$

$$p_{\varepsilon}(L_{sr}(t)) = 1 - \frac{\min\{d(L_{sr}(t), L_{sr}(c_{\varepsilon}))\}}{\sum_{\tau=0}^{e} \min\{d(L_{sr}(t), L_{sr}(c_{\tau}))\}}, \varepsilon \in [0, e]. \tag{9}$$

To perform ethnicity classification of a source DLM, we choose the corresponding subclass in the DLM structure. The center DLM of the subclass of interest has the smallest distance to the target DLM. In Eq. (8), C_i represents a subclass in the DLM structure, c_i represents the center DLM in C_i and C_{min} belongs to the set of C_i. C_{min} is the value of C_i with the minimum distance from the source DLM $L_{sr}(t)$ to the center DLM $L_{sr}(ci)$. The computation of C_{min} can be regarded as a preprocessing step for the ethnicity classification function. The classification function is based on the C_{min}. Eq. (9) shows the ethnicity classification function as a probability representation. Parameter e represents the index of ethnic backgrounds, ε represents the ethnic background for which we want to compute the probability.The p_{ε} represents the probability that the ethnic background of $L_{sr}(t)$ is ε. Each DLM in C_{min} is associated with data for an ethnic background τ. We classify the DLM into several groups by ethnic background. We compute the distance from the source DLM $L_{sr}(t)$ to DLMs in Cmin. Comparing the distances, we obtain a set of DLMs with the minimum distance to $L_{sr}(t)$ in different ethnic background groups. Using the set of DLMs, we obtain the probability values of $L_{sr}(t)$ for all ethnic backgrounds. The ethnic background with the largest value of p is the final classification result.

5. Experiments

We apply our ethnicity classification method to data from four public facial databases, Texas3D [36], FRGC2.0, BU-3DFE [37] and BU-4DFE [38]. Texas3D includes 1149 scans of 117 individuals, featuring different facial expressions and accompanied by landmark information. We choose 212 facial data of 106 persons to construct the training set and other facial data(about 900 scans) to create the testing set. FRGC2.0 is a large facial database of more than 500 individuals and 4000 scans. The facial data in FRGC2.0 are not registered and have no landmark information. The facial scans in FRGC feature various facial expressions. We select 367*3 scans of 367 individuals(the facial data in Fall2003) as the training set, and other facial data(about 2200 scans) of the database as the testing set. BU-3DFE is a facial expression database with accurate facial attribute information(7 expressions with 4 level, 6 ethnicity groups). BU-4DFE is another facial expression database with continuous sampling data. For BU-3DFE, we select 7 samples (7 expressions with level 2) from each person to construct training set and other facial data of the database as the testing set. For BU-4DFE, we select 18 samples from each person (6 expressions * 3 samples) to construct training set and other facial data of the database as the testing set. We use facial database facewarehouse to be the template facial database for DLM reconstruction.

Table 1
Ethnicity classification rates in different local coordinate systems. C1:C(11,10,17,19), C2:C(7,4,17,19), C3:C(7,4,17,18), C4:C(1,0,17,19), C5:C(15,14,20,23).

Ethnic group	C1	C2	C3	C4	C5
Asian	93.3%	91.1%	82.3%	83.3%	79.2%
Non-Asian	97.7%	95.5%	97.5%	92.3%	90.1%
Average	95.1%	94.3%	94.1%	89.2%	87.5%

Table 2
Ethnicity classification rates of two facial landmarks detection methods applied to FRGC2.0 data.

Ethnic group	Shape Regression [35]	DEST [32]	Face++
Asian	86.4%	98.3%	97.7%
Non-Asian	75.7%	93.6%	92.6%
Average	82.3%	95.4%	94.8%

5.1. Ethnicity classification of Texas3D data

The facial data in Texas3D include 25 manually selected landmarks of different facial features. We build the DLM directly from landmarks' positions. The facial data in Texas3D are not registered. The local coordinate systems used by different landmarks affect the DLM measurement. We obtain different DLM clustering results for Texas3D data by using different local coordinate systems. The local coordinate system is created from four landmarks, represented by $C(a, b, c, d)$ (with a, b, c, d representing the landmark indices). Using the landmark index, we obtain the global point coordinates $a(x, y, z)$, $b(x, y, z)$, $c(x, y, z)$ and $d(x, y, z)$. The local coordinate system is constructed from the global point coordinates.

$$x_{local} = a(x, y, z) - b(x, y, z)$$
$$y_{local} = c(x, y, z) - d(x, y, z)$$
$$z_{local} = x_{local} \times y_{local}$$

The landmarks in the local coordinate system should be robust to facial expressions. We select the landmarks around the nose area. The reason for this choice is that positions of landmarks around the nose area are less affected by facial expressions. We select different combinations of landmarks to build the local coordinate systems. The clustering results also depend on the local coordinate system being used. Fig. 4 shows the facial data together with landmark indices in Texas3D and receiver operating characteristic curve (ROC) for the ethnicity classifier based on different clustering results. Table 1 shows the classification rate. We choose the clustering result with the best ROC curve to construct the ethnicity classifier. The classification results show that the local coordinate system should be constructed based on facial landmarks around the nasal region. We select the C1 to construct local coordinate system in our classifier.

5.2. Ethnicity classification in FRGC2.0

The facial data in FRGC2.0 include various facial expressions without the accompanying facial landmarks. The point cloud of facial data is not registered. Detection of landmarks influences the classification rate. We compare the classification rate with three landmarks' sets: 3D shape regression [35] with 14 landmarks, One Millisecond Deformable Shape Tracking (DEST) [32] with 68 landmarks and Face++(business tool from Megvii: www.faceplusplus.com.cn) with 106 landmarks. Fig. 5 shows the ROC ethnicity classification results based on different landmarks' sets. It shows that the performance of classification converges to certain landmarks' set. In Table 2, we show the classification rate based on different landmarks' sets. The comparison of several ex-

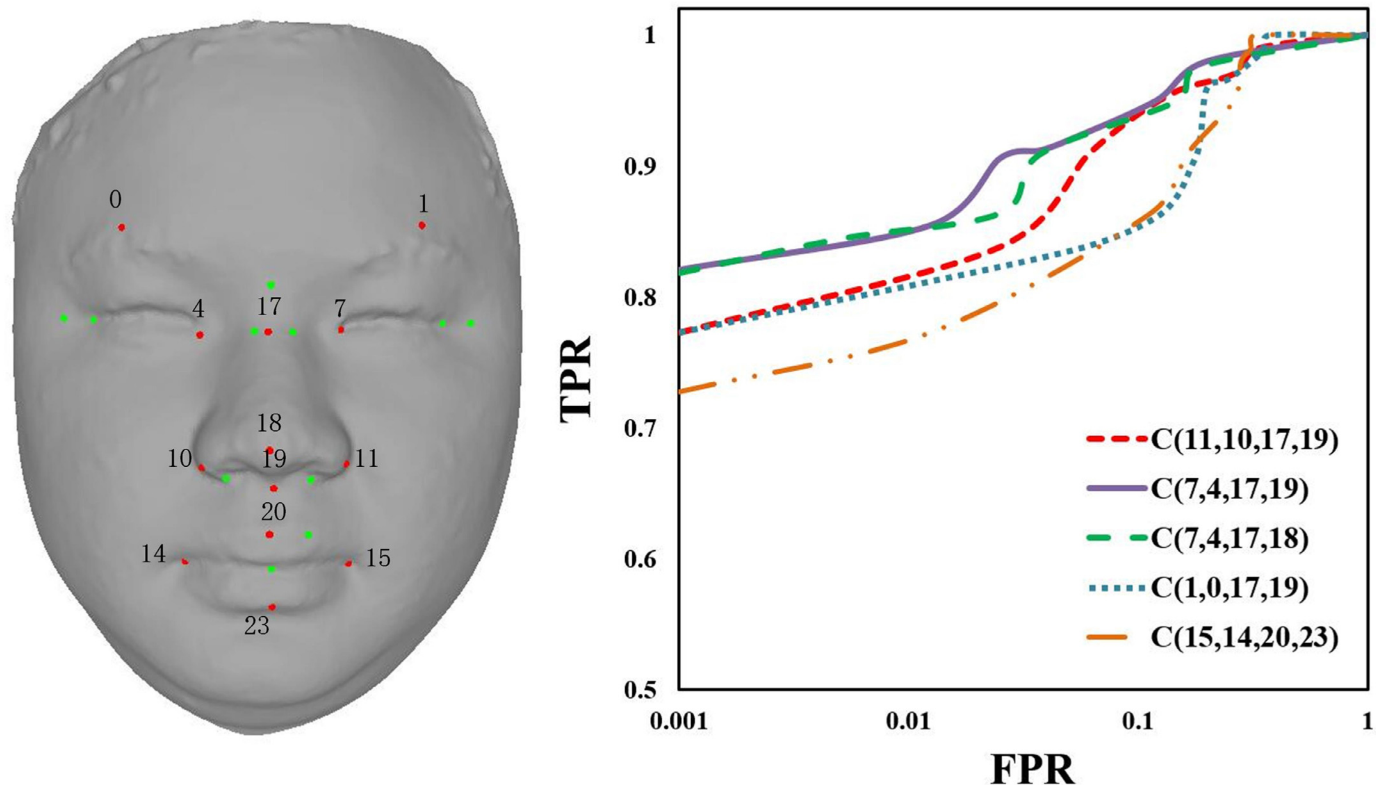

Fig. 4. The left picture is Facial landmarks with index data from Texas3D. The selected landmarks surrounding the nose area are relatively robust to facial expressions. The right picture is Ethnicity classification ROC results obtained by using different local coordinate systems for the analysis of Texas3D data.

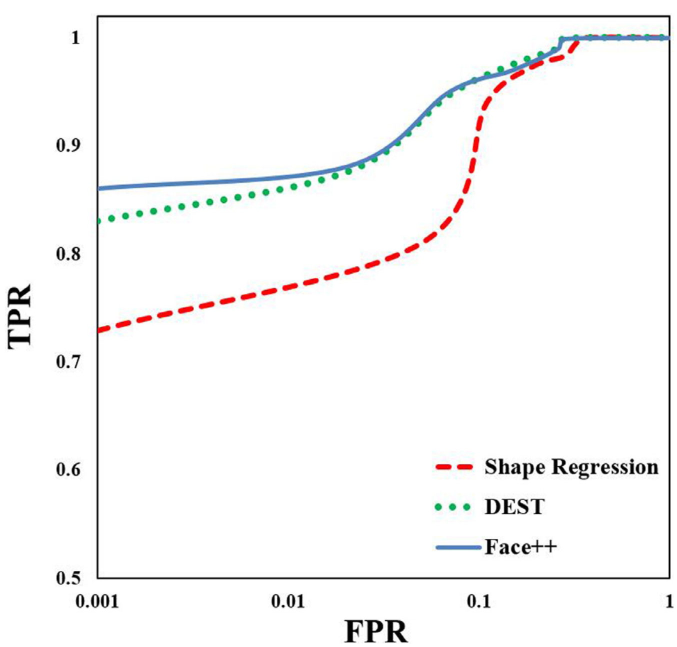

Fig. 5. Ethnicity classification ROC results of two different landmarks detection methods applied to FRGC2.0 data.

isting methods to our method is shown in Table 3. Some classification methods extract facial features from 3D face surface and 2D facial image. Some classification methods construct classifier based on 3D mesh without texture information. The result shows that our method can achieve better classification rate based on 3D landmarks positions, which is robust to different texture information.

Table 3
Ethnicity classification rate comparison of our method to earlier methods in FRGC2.0.

Method	Object	Classifier	Asian	Non-Asian	Average
Zhang [7]	3D+2D	Adaboost	87.65%	82.36%	85.62%
Ding [29]	3D+2D	AdaBoost	93.35%	89.49%	92.06%
Huang [28]	3D+2D	Adaboost	94.13%	96.90%	95.45%
Xia [21]	3D	Random Forest	92.12%	88.67%	91.60%
Our	3D	Shape Space	97.30%	93.60%	95.40%

5.3. Ethnicity classification in BU-3DFE and BU-4DFE

The facial data in BU-3DFE and BU-4DFE include various facial expressions with facial landmarks. The ethnicity labels are more accurate. In this part, we evaluate the classification performance of our method for different ethnicity groups classification tasks. The different facial attributes (expression and gender) also be considered seriously. For BU-3DFE, the facial data can be divided into 4 ethnic groups: WH/BL/AE/IML (WH:White&Middle-east Asian, BL:Black, AE:East Asian, IL:Indian&latino-Hispanic). We evaluate the ethnicity classification performance for different facial data set with certain facial attribute(expression and gender). In Table 4, we show the classification rate for certain facial data set from BU-3DFE. For BU-4DFE, the ethnicity labels do not exist for each facial data. We label the ethnicity information manually. To avoid the label errors, we delete the ethnicity classification task for IML. In Table 5, we show the classification rate for certain facial data set from BU-4DFE.

5.4. Summary and comparison

We evaluate the classification rate in different facial databases. For test in Texas3D, we evaluate the influence of local coordinate

Table 4
Ethnicity classification rate based on different groups in BU-3DFE.

Ethnic group	Facial expressions							Gender	
	Neutral	Angry	Disgust	Fear	Happy	Sad	Surprise	Female	Male
WH	0.9801	0.9501	0.8783	0.9735	0.9513	0.9113	0.9018	0.982	0.919
BL	0.9888	0.8722	0.8388	0.9222	0.9222	0.8111	0.8388	0.9171	0.944
AE	0.9787	0.9456	0.9422	0.949	0.9803	0.95	0.9343	0.9763	0.9338
IML	0.8933	0.8216	0.8416	0.8166	0.8066	0.91	0.86	–	0.89
Average	0.95	0.91	0.89	0.93	0.95	0.92	0.91	0.96	0.92

Table 5
Ethnicity classification rate based on different groups in BU-4DFE.

Ethnic group	Facial expressions						Gender	
	Angry	Disgust	Fear	Happy	Sad	Surprise	Female	Male
WH	0.9522	0.922	0.963	0.9489	0.9365	0.9321	0.971	0.931
BL	0.9122	0.913	0.9423	0.9411	0.8965	0.9211	0.923	0.946
AE	0.9389	0.934	0.944	0.9756	0.9602	0.9313	0.963	0.935
Average	0.932	0.923	0.9501	0.962	0.951	0.929	0.952	0.934

Table 6
Average ethnicity classification speed for data from several facial databases.

Facial database	Texas3D	FRGC2.0	BU-3DFE	BU-4DFE
Average time	2.2S	2.6S	1.8S	2.9S

system. The local coordinate system should be constructed by facial landmarks which is robust to facial expressions. For test in FRGC2.0, we evaluate the influence of the different facial landmarks detection in our method. The detection method based on combining 2D and 3D facial data obtain more accurate classification rate. The entire process is automated and requires no manual adjustments. We also compare different classification methods in FRGC2.0. Although such methods can achieve high ethnicity classification rates when tested on publicly available facial databases, the applications of such methods are additionally limited by the need for complex preprocessing and feature extraction. The feature extraction part of our method is simple, with the DLM measures obtained by a linear computation. For test in BU-3DFE and BU-4DFE, we evaluate the influence of different expressions and genders for accurate ethnicity classification tasks. The classification results show that our method is robust to different facial expressions and genders for accurate classification tasks. In our method, the ethnicity classifier is constructed based on facial landmarks. The computation of pre-process in our method is lower relatively. Table 6 shows the computation times of our ethnicity classification method applied to different facial databases.

6. Conclusions

We propose a novel automated ethnicity classification method using 3D facial landmarks. We construct a DLM representation of facial landmarks to model the relative positions of facial features. Subsequently, we measure DLMs in the mathematical manifold space called the Kendall shape space. The measurement of DLMs in this space is robust to rigid transformations. Using the DLMs measurement results, we cluster DLMs into different classes. The ethnicity classification equation is constructed based on DLM subclasses. The full framework does not require complex preprocessing involving mesh reconstruction and feature extraction. The landmarks are easily obtained by landmarks detection methods. The DLMs measurement in the Kendall space involves an extremely fast linear computation. Our method, without auxiliary 2D facial textures, can achieve classification rates similar to those of multi-model feature-based methods.

Declaration of Competing Interest

The authors declare that they do not have any financial or non-financial conflict of interests.

Acknowledgments

This research was partially supported by the National Key Research and Develop Program of China (No. 2017YFE0100500, No. 2017YFB1002600), the Chinese High-Technical Research Development Foundation Program (No. 2015AA020506), Beijing Natural Science Foundation of China (No. 4172033), We thank the face databases providers (FRGCv2.0, Texas3D and FaceWareHouse) and method's code provider in Github.

References

[1] G. Shakhnarovich, P. Viola, B. Moghaddam, A unified learning framework for real time face detection and classification, in: IEEE International Conference on Automatic Face and Gesture Recognition, 2002.

[2] G. Guo, G. Mu, A study of large-scale ethnicity estimation with gender and age variations, in: IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, 2010, pp. 79–86.

[3] H. Han, C. Otto, X. Liu, A. Jain, Demographic estimation from face images: human vs. machine performance, IEEE Trans. Pattern Anal. Mach. Intell. 37 (6) (2015) 1148–1161.

[4] H. Lian, B. Lu, E. Talikawa, S. Hosoi, Gender recognition using a min-max modular support vector machine, in: International Conference on Natural Computation, 2005.

[5] H. Lu, H. Lin, Gender recognition using adaboosted feature, in: IEEE International Conference on Natural Computation, 2007.

[6] H. Zhao, 3d facial gender classification based on multi-angle LBP feature, Acta Autom. Sin. 38 (9) (2012) 1544.

[7] G. Zhang, Y. Wang, Multimodal 2d and 3d facial ethnicity classification, in: Fifth International Conference on Image and Graphics, IEEE Computer Society, 2009, pp. 928–932.

[8] B. Xia, B. Amor, D. Huang, et al., Enhancing gender classification by combining 3d and 2d face modalities, in: Signal Processing Conference, IEEE, 2013, pp. 1–5.

[9] S. Hosoi, E. Takikawa, M. Kawade, Ethnicity estimation with facial images, in: IEEE International Conference on Automatic Face and Gesture Recognition, 2004, pp. 195–200.

[10] H. Han, C. Otto, X. Liu, A. Jain, Demographic estimation from face images: human vs. machine performance, IEEE Trans. Pattern Anal.Mach. Intell. 37 (2015) 1148–1161.

[11] X. Lu, A. Jain, Ethnicity identification from face images, in: SPIE International Symposium on Defense and Security: Biometric Technology for Human Identification, 2004, pp. 114–123.

[12] G. Guo, G. Mu, A study of large-scale ethnicity estimation with gender and age variations, in: IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, 2010, pp. 79–86.

[13] A. Moeini, H. Moeini, Pose-invariant gender classification based on 3d face reconstruction and synthesis from single 2d image, Electron. Lett. 51 (10) (2015) 760–762.

[14] D. Huang, H. Ding, C. Wang, et al., Local circular patterns for multi-modal facial gender and ethnicity classification, Image Vis. Comput. 32 (12) (2014) 1181–1193.
[15] S. Shan, M. Khalil-Hani, S. Radzi, R. Bakhteri, Gender classification: a convolutional neural network approach, Turk. J. Electr. Eng. Comput. Sci. 24 (3) (2016) 1248–1264.
[16] X. Han, H. Ugail, I. Palmer, Gender classification based on 3d face geometry features using SVM, in: International Conference on Cyberworlds, IEEE Computer Society, 2009, pp. 114–118.
[17] Y. Hu, J. Yan, P. Shi, A fusion-based method for 3d facial gender classification, in: International Conference on Computer and Automation, Engineering, 2010, pp. 369–372.
[18] L. Ballihi, B. Amor, M. Daoudi, A. Srivastava, D. Aboutajdine, Geometric based 3d facial gender classification, in: International Symposium on Communications Control and Signal Processing, IEEE, 2012, pp. 1–5.
[19] B. Xia, B. Amor, H. Drira, et al., Gender and 3d facial symmetry: what's the relationship? in: IEEE International Conference and Workshops on Automatic Face and Gesture Recognition, IEEE, 2013, pp. 1–6.
[20] B. Xia, B. Amor, H. Drira, M. Daoudi, L. Ballihi, Combining face averageness and symmetry for 3d-based gender classification, Pattern Recognit. 48 (3) (2015) 746–758.
[21] B. Xia, B. Amor, M. Daoudi, Joint gender, ethnicity and age estimation from 3d faces: an experimental illustration of their correlations, Image Vis. Comput. (2017).
[22] S. Gilani, F. Shafait, A. Mian, Biologically significant facial landmarks: how significant are they for gender classification, in: 2013 International Conference on Digital Image Computing: Techniques and Applications, 2013, pp. 1–8.
[23] R. Tokola, A. Mikkilineni, C. Boehnen, 3d face analysis for demographic biometrics, in: International Conference on Biometrics, IEEE, 2015, pp. 201–207.
[24] T. Cootes, C. Taylor, D. Cooper, J. Graham, Active shape models-their training and application, Comput. Vis. Image Understanding 61 (1) (1995) 38–59.
[25] T. Cootes, G. Edwards, C. Taylor, Active appearance models, IEEE Trans. Pattern Anal. Mach. Intell. 23 (2001) 681–685.
[26] P. Perakis, G. Passalis, T. Theoharis, I. Kakadiaris, 3d facial landmark detection under large yaw and expression variations, IEEE Trans. Pattern Anal. Mach. Intell. 35 (2013) 1552–1564.
[27] B.J. Frey, D. Dueck, Clustering by passing messages between data points, Science 315 (5814) (2007) 972.
[28] D. Huang, H. Ding, C. Wang, et al., Local circular patterns for multi-modal facial gender and ethnicity classification, Image Vis. Comput. 32 (12) (2014) 1181–1193.
[29] H. Ding, D. Huang, Y. Wang, et al., Facial ethnicity classification based on boosted local texture and shape descriptions, in: IEEE International Conference and Workshops on Automatic Face and Gesture Recognition, IEEE, 2013, pp. 1–6.
[30] D. Kendall, A survey of the statistical theory of shape, Stat Sci. 4 (2) (1989) 87–99.
[31] D. Kendall, Shape manifolds, procrustean metrics, and complex projective spaces, Bull. Lond. Math. Soc. 16 (2) (1984) 81–121. Landmarks detection
[32] V. Kazemi, J. Sullivan, One millisecond face alignment with an ensemble of regression trees, in: Computer Vision and Pattern Recognition, IEEE, 2014, pp. 1867–1874.
[33] M. Segundo, L. Silva, O. Bellon, C. Queirolo, Automatic face segmentation and facial landmark detection in range images, IEEE Transactions on Systems, Man, and Cybernetics, 2010.
[34] Y. Wang, J. Liu, X. Tang, Robust 3d face recognition by local shape difference boosting, IEEE Trans. Pattern Anal. Mach. Intell. 32 (2010) 1858–1870.
[35] F. Sukno, J. Waddington, P. Whelan, 3-d facial landmark localization with asymmetry patterns and shape regression from incomplete local features, IEEE Trans. Cybern. 45 (9) (2015) 1717.
[36] S. Gupta, K. Castleman, M. Markey, et al., Texas 3d face recognition database, in: IEEE Southwest Symposium on Image Analysis and Interpretation, 2010, pp. 97–100.
[37] L. Yin, X. Wei, Y. Sun, et al., A 3d facial expression database for facial behavior research, in: International Conference on Automatic Face and Gesture Recognition, 2006, pp. 211–216.
[38] L. Yin, X. Chen, Y. Sun, et al., A high-resolution 3d dynamic facial expression database, in: International Conference on Automatic Face and Gesture Recognition, 2008, pp. 1–6.