Multimed Tools Appl
https://doi.org/10.1007/s11042-018-6018-1

Multi-ethnical Chinese facial characterization and analysis

Cunrui Wang[1,2] · Qingling Zhang[1] · Xiaodong Duan[2] · Jianhou Gan[3]

Received: 10 October 2017 / Revised: 5 April 2018 / Accepted: 16 April 2018

Abstract Facial image based characterization and analysis of ethnicity, which is an important index of human demography, have become increasingly popular in the research areas of pattern recognition, computer vision, and machine learning. Many applications, such as face recognition and facial expression recognition, are affected by ethnicity information of individuals. In this study, we first create a human face database, which focuses on human ethnicity information and includes individuals from eight ethnic groups in China. This dataset can be used to conduct psychological experiments or evaluate the performance of computational algorithms. To evaluate the usefulness of this created dataset, some critical landmarks of these face images are detected and three types of features are extracted as ethnicity representations. Next, the ethnicity manifolds are learnt to demonstrate the discriminative power of the extracted features. Finally, ethnicity classifications with different popular classifiers are conducted on the constructed database, and the results indicate the effectiveness of the proposed features.

Keywords Chinese ethnicity · Manifold learning · Ethnicity classification

✉ Xiaodong Duan
duanxd@dlnu.edu.cn

Cunrui Wang
curui@mail.neu.edu.cn

Qingling Zhang
qlzhang@mail.neu.edu.cn

Jianhou Gan
ganjh@ynnu.edu.cn

1 Institute of System Science, Northeastern University, Shenyang, China

2 Dalian Key Lab of Digital Technology for National Culture, Dalian Nationalities University, Dalian, China

3 Key Laboratory of Educational Informatization for Nationalities, Yunnan Normal University, 650500, Kunming, China

Published online: 23 May 2018

1 Introduction

The human face conveys the most direct, effective, and essential information for daily communication and interaction. The human facial area contains various kinds of semantic information, such as ethnicity, age, gender and expressions and so on. As illustrated in Fig. 1, this semantic information is processed rapidly and sequentially when a face is viewed. The authors of, [11] stated that the recognition of racial and ethnic groups is a top priority when human faces are seen and recognized.

The population of China consists of multiple-ethnicities, and their mutual influence and fusion have shaped Chinese history, traditions and culture. The diverse facial features that have emerged among different ethnicities because the geographic isolations and genetics, which have been investigated by experts in anthropology for decades [4, 9, 45]. The facial variations of ethnic groups [20, 21] in China have been studied experimentally in anthropometry and those results have been used to analyze the origin, evolution and fusion of ethnic groups [35, 38, 39].

Recently, ethnic representation and classification have become increasingly popular in the computer vision community, and it is well known that the performance of computational algorithms is affected by ethnic facial features. Due to the variations in geometric shape and texture of the faces from different cultures and ethnicities, advancements in face recognition, and ethnicity classification rely heavily on databases of faces from different ethnicities. In addition, advancements in 2D and 3D techniques in morphometry and anthropometry enables us to measure some non-linear features, such as the angle and curvature of an irregular shape. These developments provide opportunities for us to investigate ethnicity-related facial issues, such as multi-ethnic face recognition, facial expression recognition and ethnic classification and identification. For example, many algorithms have been proposed for separating Asians and non-Asians, including boosting algorithms [8, 37], Principal Component Analysis (PCA) [7, 25], and linear discrimination analysis (LDA) [23]. Meanwhile, the classification of Asians, Africans and Caucasians is investigated based on facial features using an SVM classifier in [18, 22]. The authors of, [28, 29, 40] studied the characteristics of the eastern and western faces using virtual vision coding techniques. The authors of, [14, 15] have trained an RBF neural network to recognize the gender and race of the faces in the FERET database [27]. Moreover, the authors of [39] have examined the skin textures of the 56 ethnicities in China, This work provides substantial evidences and supports for the research on ethnic origin and fusion. For example, the skin textures of individuals from Tibetan and Qiang show the northern China group characteristics, which reveals that the origin of Tibetan people is highly related to that of Qiang people. In addition, the skin textures of individuals from Han demonstrate obvious multi-ethnic characteristics, which is an evidence of ethnic fusion in the long history of China.

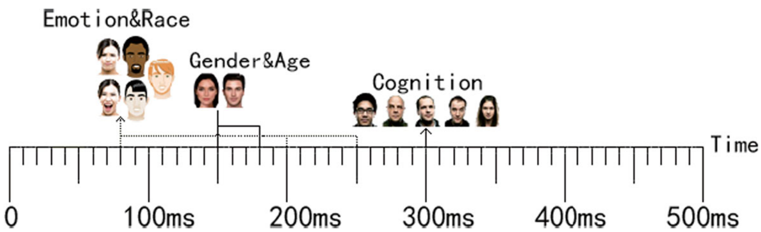

Fig. 1 The scheme of human facial information recognition order

The experiments and evaluations from both anthropometric and computational researches show the desperate need for a multi-ethnic Chinese face database of sufficient size and demonstration of its usefulness. In anthropometric studies, the researchers in [35, 38, 39] try to find the biological variations among different Chinese ethnics with the aim of providing the evidences for ethnicity definition and recognition. However, these results are obtained mainly based on experts' experience and such investigations are usually insufficient, due to lack of effective face data collection and limitations of the empirical features that are used in the study [38, 39]. Hence, it is necessary for us to construct a Chinese multi-ethnic face database for the study of anthropometric and face-related computer vision algorithms [43]. This database should contain face images of individuals from different ethnic groups in China. To conduct a full-scale and insightful investigation of Chinese ethnics facial differences, a facial database of individuals of different ethnicities is desperately required by the research community. In this paper, our first task is to collect faces of individuals of eight important ethnicities in China, which can be used to extract the facial geometric and texture features with the purpose of description and representing different ethnicities.

To enable sufficient and effective anthropometric and computational analysis, another task in this paper is to define a set of features for facial ethnic descriptions. Unlike previous work in anthropometric study, we define and select the discriminative features based on automatically detected landmarks. The face landmarks represent the most important face components' location, and are widely used in face image based processing and recognition tasks. For example, a carefully designed landmark detector [19] is used to detect face contours, which are then used as salient edges to guide the image deconvolution for deblurring. A active shape model based facial landmarks detector is adopted in this work, since the images in the constructed database contain only frontal faces. The collected data are then evaluated based on the selected characterizing features by conducting ethnic manifold learning and classification. In summary, we make several contributions in this paper. First, a Chinese multi-ethnic face database, which contains eight ethnicities, is constructed. The frontal face images of each ethnicity are digitalized and labeled. Second, three types of features, namely, distance, angular and ratio features, are defined based on the automatically detected landmarks. The most discriminative features are extracted by a data-driven method for the purpose of ethnic representation. Finally, face images and the selected features are evaluated by performing manifold learning and ethnicity classification. Some benchmark results are obtained for further study in the future.

The structure of this paper is as follow: Section 2 gives a brief summary of the existing work on face ethnic databases. Section 3 describes the data collection process for the constructed database. In Section 4, the details of landmark detection, and feature extraction are presented. We evaluate the database for ethnicity classification in Section 5 and conclude our study in Section 6.

2 Related work

Ever since the first study in anthropology, the ethnicity-related data have been collected for various research purposes. Such existing data ranges from body to face and focuses on differences among ethnic groups. The first 3D human body database [30] was collected by Civilian American and European Surface Anthropometry Resource (CAESAR), in which the northern American, Netherlands and Italy are the main areas that were considered. The Texas 3D face database [13] is a popular face database, which contains 3D faces from the main races from around the world, such as Caucasian, African, Asian, East-Indian and

Latino. Another database [31, 44] was collected by National Institute of Occupational Safety and Health (NIOSH), which includes geographic information and traditional anthropometric measurements, such as human body structure, race, gender and age.

The collection of Chinese face datasets began decades ago, when several databases were created for different purposes. The CAS-PEAL face database [12], was created for research on face recognition. The faces in this database were collected from participants with different poses and expressions under different illumination conditions. The CUN database [10] is the first Chinese face database, which contains the faces of subjects from 56 ethnic groups in China. Variations of illumination, pose, background, and expressions are recorded in this database. Recently, the advancements in 3D imaging techniques enabled the collection of the 3D human faces. The first 3D Chinese face database, which is called SizeChina [2], contains high-resolution 3D faces of 2000 adults. The research based on this database has identified subtle facial differences between Chinese and Caucasian individuals.

All of these databases were collected for different purposes, but none of them focuses on the ethnic description and recognition. This is the reason why this work attempts to collect face data of different ethnic groups in China. Although the databases that are mentioned above contain facial or body information of individuals of different ethnicities, they are not focused on ethnic representation and classification. More specifically, these existing datasets may not suitable for computational algorithm development or evaluation, due to either lack of effective data for specific ethnic groups or failing to provide discriminative ethnic features. This is why we construct a Chinese multi-ethnic face database in this paper, which contains effective facial images for characterizing features for ethnicity representation and classification.

3 Data collection

3.1 Background description

China is a large country, not only geographically but also demographically. The population of China is consists of one major ethnic group (Han) and many minorities, such as Zhuang, Mongolian, Manchu and Uyghur. The population sizes of some ethnic groups are relatively small, which causes the difficulty in data collection. As shown in Table 1, seven ethnicities are considered in this paper, with ethnic Han be separated into Northern Han and Southern Han geographically. Hence, eight ethnic groups are considered in our database and the following experiments. We start with the ethnicities that have a population size of over one

Table 1 The demographic information on the ethnicities in China

Ethnicity	Population size	Proportion
Han	1,220,844,520	91.6399%
Zhuang	16,926,381	1.2700%
Manchu	10,387,958	0.7794%
Uygur	10,069,346	0.7555%
Tibetan	6,282,187	0.4713%
Mongolian	5,981,840	0.4488%
Korean	1,830,929	0.1374%

million. To guarantee the genetic purity and ethnic uniqueness of the collected data, the participants of each ethnic group are selected carefully. It is required that the participants of the same ethnic group be from the same living location and be third-generation members of the same group, that is to say, the participant's parents and grand parents must also belong to the same ethnic group. Students of different ethnics in China are the main participants in the facial data collection.

Long-term location isolation results in genetic and cultural isolation, which may cause many biological differences including facial feature differences. Inter-group differences are similar to differences caused by race. Moreover, the geographic isolation will also result in intra-group differences, such as the facial differences between the Northern Han and Southern Han peoples. This is the reason why we distinguish Northern Han subjects from Southern ones in our database.

The data collection started in 2012. We set up a digital facial image capturing system with soft lighting equipment. The facial images of 500 participants from Korean, Zhuang, Uygur, Mongolian, and Tibetan (collected in 2012) were collected. Each ethnicity has 100 individuals and the resolution of the captured images is 400*600 pixels. In 2016, the facial data of Manchu and Han ethnic groups were added into the database. The Han ethnicity is divided into two ethnic groups: Northern Han ethnic group and Southern Han ethnic group. Therefore, the database contains 7 ethnic groups, with Han ethnicity being split into two sub-classes i.e. Northern Han and Southern Han. Finally, there are eight groups of data in the collected facial image database: Northern Han, Southern Han, Korean, Man, Mongolian, Uygur, Tibetan, and Zhuang. Each group includes 200 facial images, 100 males and females. A frontal image is taken from each individual. Finally, the database contains 1600 facial images in total. The age of the participants ranges from 18 to 22. It is shown in [5] that participants of this age range can sufficiently represent the features differences that are caused by cultural differences.

3.2 Data collection setup

All of the images are collected in a digital photography studio, in which facial images of individuals ethnicities are captured. As illustrated in Fig. 2, three cameras are setup and slaved by one computer is used to capture the individual's frontal images and profile images

Fig. 2 Configuration of the digital photography studio

(left and right). To normalize the collected images afterwards, the photography studio is calibrated with a scale.

To eliminate impact of illumination, the lighting conditions during the collection are strictly controlled. Together with multiple ceiling lights, four project lights and three reflectors are used to tune the illumination environment. The participants' facial images are captured without makeups, and no expression is conveyed.

3.3 Data collection process

The participants are selected from the undergraduate students in Dalian Minzu university. The chosen individuals of each ethnic group grew up in same locations, and their families have belonged to this ethnic group for at least three generations. For example, for a student to be chosen for the collection of ethnic Mongolian facial images, it is required that his grandfather/grandmother and parents must be from ethnic Mongolian. Furthermore, all the selected participants of the same ethnic group must live in the same location. This requirement could guarantees the genetic purity of participants and the ethnic uniqueness of the facial features, which are essential for ethnic representation.

The data collection process lasts approximately 6 months, since we capture each participant at least 3 times. For each ethnic group, 200 participants are selected as the final data source. Their face images are captured and digitalized into RGB images with a resolution of 480×640. The collected images are assigned with ethnic labels based on each participant's self-rated ethnic information. Figure 3 shows some sample images from the collected database. The database is free for public research, which could be download by the link: http://zs.dlnu.edu.cn/7minzu8ethnicGroups.rar.

4 Landmark and feature extraction

To evaluate the usefulness of this constructed dataset, landmark detection [19, 42] is the first step of pre-processing before feature extraction. The real-time detection method STASM [3, 24] is adopted to detect 77 facial landmarks for feature extraction. As shown in Fig. 4, the landmarks are located near important facial components, such as eyes, nose, mouth and facial contour line, which will encode its ethnicity information. The detection can be finished in real time. It takes only approximately 30ms to detect 77 landmarks on an image of resolution 480×640.

The following feature extraction approach is based on the detected landmarks. Detection accuracy is very important for the ethnicity representation. To evaluate the accuracy of the STASM method, 100 images are selected randomly from the constructed database. Then, 5 landmarks are manually labeled on the selected images, including the two mouth corners, nose tip, right-eye inner corner and left-eye outer corner. The manually identified landmarks serve as the ground truth, based on which the detection accuracy of the STASM method is evaluated. The pixel distances of the detected landmarks from the ground truth are reported as the detection errors. The accuracies of the detection are illustrated in Fig. 4. It can be seen that 90% of the detection errors are less than 5 pixels, which means that STASM detects eye corners (landmarks #31 and #45) and nose tip (landmark #53) quite well. The mouth corners (landmarks #60 and #56) are difficult to detect, as only 60% of the errors are less than 5 pixels and 90% of the errors are less than 10 pixels. The muscle around the mouth area are much more morphable than those in other areas of the face, which increases the detection difficulty.

Fig. 3 Sample images and the detected landmarks of the collected database

Based on the detected landmarks, three types of the features are defined: namely distance, angular and ratio features. For the distance features, the Euclidean distance between any pair of the landmarks is calculated and recorded. For the angular features, any three landmarks are used to form a triangle, and three interior angles are then calculated. The degrees of the

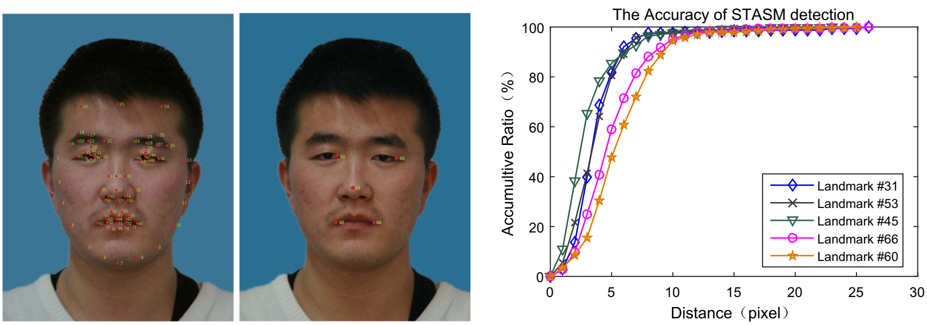

Fig. 4 The face landmarks and detection error curve of the 5 selected points

interior angles are represented in radians and used as angular features. For the ratio features, any two distances are used to calculate the proportion. The features are normalized first and then used to form a feature vector, which servers as a facial representation. The details are given as follows.

4.1 Ethnicity representation

Once the facial landmarks have been detected accurately, the ethnicity information that is contained in the facial images can be extracted and represented based on multiple geometric features. Given all 77 facial landmarks, three types of geometric features are constructed, namely distance, angular and ratio features. In this section, we will describe in detail the extraction and representation of these features.

(a) Distance features: The Euclidean distance between each pair of the landmarks is calculated and the distances between all pairs are compiled into a single feature vector. The facial structures of individuals of different ethnicities vary substantially, which is reflected in the distance between pairs of landmarks accordingly. Given two landmarks $A(x_i, y_i)$ and $B(x_j, y_j)$ on a face image, the distance between them can be obtained as follows:

$$d(A, B) = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}, 1 <= i \neq j <= N, \tag{1}$$

where N is the total number of landmarks. For all pairs of landmarks, the distances are encoded in distance-feature vector $D = (d_1, d_2, ..., d_n)$, and the dimension of the vector is $n = C_{77}^2 = 2926$. The left picture of Fig. 5 gives some samples of distance feature.

(b) Ratio features: The statistics of facial features from anthropological research [16, 36] suggests that the ratios between different distances vary from one ethnicity to another. Hence, for every pair of the distances, a ratio is calculated to represent facial ethnicity as follows:

$$r(d_i, d_j) = \frac{d_i}{d_j}, 1 <= i \neq j <= n, \tag{2}$$

where n is the number of the distance features. The ratios between all pairs of distances are encoded in ratio-feature vector $R = (r_1, r_2, ..., r_n)$, which is dimension of $p = C_n^2 = C_{2926}^2 = 4,279,275$.

(c) Angular features: As shown in the right picture of Fig. 5, given any three non-collinear landmarks $A(x_i, y_i)$, $B(x_j, y_j)$ and $C(x_k, y_k)$ on a face image, a triangle ΔABC can be formed. The three interior angle of the triangle are used to encode the relative locations which reflect the facial variations among different ethnicities. Suppose the lengths of the triangle sides are a, b and c, The angle can be measured and calculated as follows:

$$\begin{cases} \alpha = arccos\frac{b^2+c^2-a^2}{2bc}, \\ \beta = arccos\frac{a^2+c^2-b^2}{2ac}, \\ \gamma = arccos\frac{a^2+b^2-c^2}{2ab}, \end{cases} \tag{3}$$

Fig. 5 The examples of distance and angular features

For all the triangles that are formed by landmarks, the interior angles are compiled into angular feature vector $G = (\alpha_1, \beta_1, \gamma_1, \alpha_2, \beta_2, \gamma_2, ..., \alpha_m, \beta_m, \gamma_m)$, which is of dimension $3m = 3 \times C_{77}^{3} = 219,450$.

4.2 Ethnicity feature selection

The features that were described in the previous section are of extremely high dimension, especially for the ratio features, thus they are inconvenient for later feature extraction and

Table 2 The anthropometric index of human frontal face

Index	Definition (Landmark numbered in Fig 4)
Breadth-height index of head	d[mid(2,12),15]/d(1,13)
Transverse frontoparietal index	d(21,30)/d(18,25)/d(1,13)
Transverse cephalo-facial index	d(2,12)/d(1,13)
Morphological facial index	d[7,mid(31,41)]/d(2,12)
Morphological upper facial index	d[mid(31,41),68]/d(2,12)
Physiognomic facial index	d(7,15)/d(2,12)
Zygomatic Mandibular index	d(4,10)/d(2,12)
Zygomatic frontal index	d(20,27)/d(2,12)
Physiognomic upper facial index	d[mid(31,41),mid(68,71)]/d(2,12)
Fronto-facial index	d[15,mid(31,41)]/d(7,15)
Physiognomic upper facial height index	d[15,mid(68,71)]/d(7,15)
Vertical cephalo-facial index	d[7,mid(31,41)]/d(53,15)
Nasal index/height-breadth index of nose	d(55,59)/d[mid(31,41),57]
Breadth-deth index of nose	d(57,53)/d(55,59)
Lip index/oral index	d(63,75)/d(60,66)

classification. Hence, an effective feature selection method should be applied to identify the most discriminant features for ethnicity representation and analysis. The necessity of feature selection lies in two aspects. First, some of the features are highly correlated and redundant. Take the distance feature as an example, if the distance between the inner eye corner and nose tip is large, the distance between outer eye corner and nose tip will also be large. Some of these features have higher discriminant power than others. Therefore, it is necessary to select the most important features, and use them as the ethnicity representation rather than all features. Furthermore, substantial work [32–34] has been done on the selection of important features for representing human faces of different races, ethnicities, and genders in anthropometry. Table 2 lists the popular anthropometric indices of facial areas. These indices are defined based on distance-features, which are illustrated in Fig. 6a. These distances are defined either horizontally or vertically, and can only measure the widths or heights of facial components that are chosen based on human experiences. Almost all of these measurements rely on the manually selected or defined landmarks, and only lengths are calculated to represent the variations in ethnicity. Moreover, these indices can not reflect the relative positions of facial components, which also contains the ethnic clues. There is no theoretical or experimental proof to guarantee that the manually selected or defined features are the best features for ethnicity representation.

In this paper, we will propose a data-driven method for selecting the ethnically significant features using the framework of the *minimal-redundancy-maximal-relevance* (mRMR) [26]. Actually, identifying the most characteristic features of the observed data is crucial in pattern recognition problems. Given a feature selection problem, the existing methods are mainly classified into two categories: filter-based methods and wrapper-based methods. In filter-based methods, individual features or subsets are evaluated according to a given criterion, which is independent of the learning algorithms [6]. In contrast in wrapper-based methods, the candidate features are evaluated via wrapping with a learning algorithm [41].

The main concept of feature selection in mRMR is that a simple combination of individually good features does not necessarily lead to good recognition performance. In other words, "the *m* best features are not the best *m* features" [26]. To identify the optimal feature set for ethnicity representation and recognition, we adopt the framework of the *minimal-redundancy-maximal-relevance* (mRMR) [26], which is a wrapper-based method. This version of mRMR involves a two-stage selection algorithm.

Fig. 6 Anthropometric index features vs. top 10 selected features. **a** The predefined anthropometric index. **b** The top 10 distance features. **c** The top 10 angular features. **d** The top 10 ratio features

First, the mRMR criterion is used to select mutually exclusive features $S = \{x_1, ..., x_m\}$ that jointly have the largest characterization power on each ethnicity class c:

$$\begin{cases} max\,\Phi(D, R) = D(S, c) - R(S), \\ D(S, c) = \frac{1}{|S|} \sum_{x_i \in S} I(x_i; c), \\ R(S) = \frac{1}{|S|^2} \sum_{x_i, x_j \in S} I(x_i, x_j), \end{cases} \tag{4}$$

where $I(x_i; c)$ is the mutual information value between individual feature and class, $I(x_i, x_j)$ is the mutual information value between two features.

When candidate features are selected, the next task is to determine the optimal number of features m. A wrapper that tests features with a classifier is utilized to determine the size of the feature set, with direct goal of minimizing the recognition error of the specific classifier on the training set. The optimal feature set can be obtained by a forward algorithm, which starts with an empty set with a specific purpose, such as ethnicity recognition. Suppose the current feature set is S, The forward algorithm will add one feature f_i into $S' = \{S, f_i\}$ if the recognition error based on S' is lower than that based on S. This process will continue until the recognition error stops decreasing. Then, the final feature set S^* is considered as the optimal features set for ethnicity representation.

Forward mRMR feature selection is conducted on distance-, ratio- and angular features individually, and 100 distance-features, 500 ratio-features and 500 angular features are selected for extracting the most discriminative features for the eight ethnicities in the constructed database. The top 10 distance-, ratio- and angular features are listed in Tables 3, 4 and 5 respectively. Since these features are all defined based on landmarks, we illustrate the selected features with the detected landmarks in Fig. 6b, d and d. which shows that the top ten features that were selected based on the data-driven method are quite different from the predefined anthropometric features. The previously selected features are often based on the local shapes and relative positions of facial components, such as the eye, nose and mouth. This approach does share several common measurements with the features that were obtained by anthropological study, such as the position and shape of the eyebrow, which suggests that the features in anthropometric measurements reflect the ethnic information. Moreover, the proposed selection methods also discover some new features that have not been identified previously by anthropologists. This finding can be considered a

Table 3 Details of selected top 10 distance-features

Feature rank	Feature details
1	d(42,44)
2	d(17,36)
3	d(52,58)
4	d(27,28)
5	d(2,58)
6	d(31,35)
7	d(32,34)
8	d(62,76)
9	d(18,22)
10	d(29,45)

Table 4 Details of selected top 10 ratio-features

Feature rank	Feature details
1	d(49,70)/d(43,76)
2	d(21,30)/d(27,43)
3	d(37,47)/d(36,48)
4	d(49,50)/d(21,36)
5	d(28,29)/d(28,43)
6	d(65,66)/d(18,33)
7	d(51,68)/d(5,33)
8	d(44,45)/d(33,37)
9	d(18,37)/d(23,76)
10	d(43,52)/d(14,62)

complementary result for the research of anthropometry in ethnicity definition and classification. More detailed feature comparisons and evaluations will be performed in Section 5 by manifold learning and ethnicity classification.

5 Evaluation of the constructed database

To evaluate the usefulness of the constructed database, two types of experiments on ethnicity manifold learning and classification are conducted in this section. First, two common manifolds are learned for the eight ethnicities in the proposed database, to evaluate the precision and effectiveness of the selected features on the collected data. The structures of the learned manifolds also show the high non-linearity and complexity of the ethnicity distribution. Ethnicity classification is also carried out to obtain the baseline results using the database, which sheds light on the usability of the database. The results show that ethnicity recognition remains a difficult task, which requires additional efforts in data collection and ethnicity feature learning.

Table 5 Landmarks and triangles related to the selected top 10 angular-features

Feature rank	Feature details	Landmarks
1	$\angle 77$	(11, 68, 77)
2	$\angle 70$	(4, 68, 70)
3	$\angle 71$	(14, 68, 71)
4	$\angle 21$	(21, 32, 37)
5	$\angle 68$	(24, 68, 71)
6	$\angle 71$	(41, 68, 71)
7	$\angle 71$	(60, 68, 71)
8	$\angle 71$	(66, 68, 71)
9	$\angle 70$	(68, 70, 71)
10	$\angle 68$	(68, 71, 73)

In the constructed database, the images of individuals of 8 ethnicities are selected for the proposed experiments. For each ethnicity, 200 images are selected and compiled to form a gallery set. The landmarks of each images are detected for extracting geometric features, which are then used to conduct ethnicity manifold learning and classification.

5.1 Manifold learning

The unique ethnic features should possess discriminant power. To evaluate the effectiveness of the selected features, the extracted distance, angular and ratio features are used to learn the ethnicity manifold structure, which could reveal the differences among ethnicities. ISOMAP [1] is a classical manifold learning approach that preserves geodesic distance in nonlinear data sets. It can preserve the data structure of an ethnic group when learning the manifold in the high dimensional space. In addition, the Laplacian manifold [17] has been proven to provide the optimal linear approximations to the eigenfunctions of the Laplace Beltrami operator on the face manifold, which facilitates investigation of the essential face manifold structures. Hence, ISOMAP and Laplacian manifolds are utilized in the proposed facial ethnical manifold learning approach.

As shown in Fig. 7, the Laplacian manifold and ISOMAP manifold are learned based on three types of features. Figure 7a and d are the manifolds that are learnt based on the distance features. It is very difficult to identify differences among the 8 ethnicities, since the data are mixed together if the distance features are used to represent the samples. In other words, the distance features are not effective for ethnicity representation. They lack discriminative power even if there are only 8 ethnicities are considered. Figure 7c and f are the ISOMAP and Laplacian manifolds that are learnt based on ration features. Some of the ethnicities are separated quite well, such as the Korean and Mongolian, especially on the Laplacian manifold. However, most of the ethnicities are still mixed with one another. Fortunately, the ISOMAP and Laplacian manifolds are also learned based on the angular features, as shown in Fig. 7b and e. In general, the eight ethnicities are separated well and some interesting correlations are identified among the easily-confused ethnicities. The Southern Han samples scatter into two clusters, which renders the Southern Han mixed with Zhuang. On both the ISOMAP and Laplacian manifolds that most of the Southern Han samples are gathered compactly into one cluster, while the rest are scattered among the Zhuang samples. Coincidentally, the ethnic Zhuang lives in the southern area of China, which overlaps with the area in which the Southern Han live. This implies that the environment has quite a significant influence on the facial features of individuals of different etnicities.

The angular features possess the best discriminative power for the eight ethnicities, since the faces from the same ethnicity cluster compactly and are mutually well separated. The angular features can well represent the samples of different ethnicities. Furthermore, the angular features that are based manifold learning can reflect the geographical distribution of the ethnicities. In other words, the ethnic groups from same or nearby areas will be near one another in angular features based subspaces. It can be seen from Fig. 7 that different ethnicities of same race has independent manifold structure, which shows the variations of facial features. The difference between Northern Han and Southern Han is not significant and both groups are absolutely different with the rest ethnicities. The samples of Zhuang, Korean, and Manchu form clear manifold structures too. However, the samples of Uygur, Tibetan, and Mongolian always mix with samples from the rest ethnicities, and fail to form an independent manifold structures.

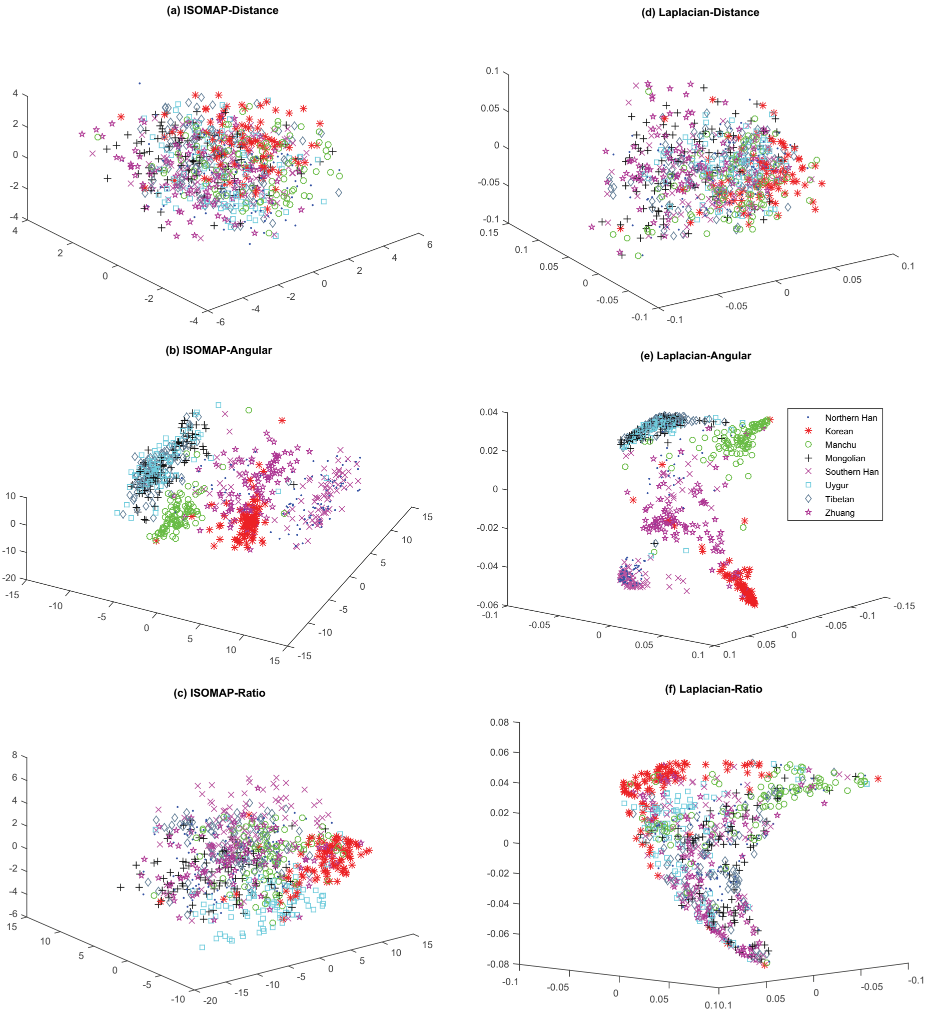

Fig. 7 Manifold structure of the ethnicities in the constructed database

5.2 Ethnicity classification

One of the purposes of constructing this database is for it to serve as a benchmark ethnic database for anthropological and computational research. In this section, we conduct ethnicity classification based on the constructed database, in which three types of features and anthropometric features are used for classification. The experiments are carried out on 800 images of all the ethnical groups, and the results are obtained by 10-fold cross validation. As recorded in Table 6, the averaged classification rates of several popular classifiers are obtained based on the collected images of the constructed database, using distance, angular and ratio features. The anthropometric features are also utilized to conduct ethnicity classification on the constructed database, and the results are compared with those of the distance, angular and ratio features, which can be extracted automatically based on

Table 6 Averaged ethnicity classification rates based on distance, angular and ratio features

Classifier	length-M	length-F	angle-M	angle-F	Ratio-M	Ratio-F
Bayesian Net	87.5	82.4	90.3	94.9	69.0	65.5
Naive Bayesian	67.9	65.5	81.3	87.8	69.8	68.0
SMO	99.8	99.6	98.8	98.6	99.8	100.0
J4.8	46.1	42.1	87.4	86.0	60.0	60.5
RBF Network	65.0	65.0	79.8	83.9	72.0	71.0
LibSVM	80.4	76.3	86.4	92.0	90.3	89.5

The ISOMAP manifolds and the Laplacian manifolds are learned based on distance, angular and ratio features respectively. The distance feature based manifolds are illustrated in figure (a) and (d), and no semantic manifold structure is formed, since the samples of all the ethnicities mixed together. The angular feature based manifolds are shown in figure (b) and (e). The ISOMAP manifold structures of Zhuang, Korean, and Manchu are well form based on angular features. However, the samples of Uygur, Mongolian, and Tibetan mix together. The samples of Northern Han also mix with the samples from Southern Han, which means that they are from same ethnicity. The manifold structures in figure (e) and figure (b) suggest that both of the them reflect the truth distribution of these ethnicities. Figure (c) shows the ISOMAP manifold based on ratio features, in which the samples of Korean, Uygur, and Southern Han form clear structures respectively. The samples of the rest ethnicities mix with each other, especially the samples of Northern Han scatters over the whole manifold space. In figure (f), the Laplacian manifold of Korean and Manchu could be separated, with the samples of rest ethnicities mixing together. The experiments are conducted based on Male and Female samples separately. The length-M means the results are based on the length feature of Male samples. Similarly, the length-F means the results are based on the length feature of Female samples

detected landmarks. Figure 8 illustrates the classification performances that are achieved based on the mentioned four specified features. The automatic features, i.e. distance, angular and ratio features outperform the anthropometric features quite significantly. Among the automatic features, the angular features can achieve the best classification performance, which is in accordance with the manifold learning results. These results suggest that the distance-based features are not sufficiently discriminative for ethnicity representation and classification. In contrast, the angular features of the human face are more characteristic for ethnic information extraction.

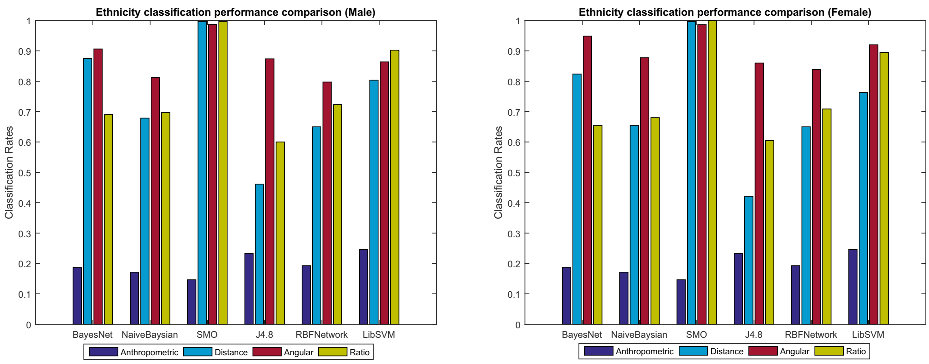

Fig. 8 Comparison of the ethnicity classification in the constructed database

6 Conclusions

Both anthropometric and computational research indicate the need for a multi-ethnic database for experiments and evaluations. This paper constructs a Chinese multi-ethnic face database, which collects 1600 face images of individuals of eight ethnicities in China. The database contains 200 images of different individuals for each ethnicity. The landmarks are detected using the STASM algorithm, and three types of features, namely,distance, angular and ratio, are forged based on detected landmarks. An mRMR-based features selection process is applied on the three types of features and the most discriminative features are selected to represent the ethnic images. Then, the ethnic manifolds are learned based on these three types of features and the ethnicity classification is conducted based on the selected features. The selected angular features, rather than distance-features, possess the most discriminative power for ethnicity representation. This work complements the related anthropometric research. The angular features can also be used to analyze the geographical distribution of ethnicities.

Acknowledgements This work is sponsored by Natural Science Foundation of China (61370146, 61672132), National Science and Technology Support Program (2013BAJ07B02) and Science & Technology Project of Liaoning Province (No.2013405003).

References

1. Balasubramanian M, Schwartz EL (2002) The isomap algorithm and topological stability
2. Ball R (2011) SizeChina: a 3D anthropometric survey of the Chinese head. TU Delft, Delft University of Technology
3. Benlamoudi A, Samai D, Ouafi A, Taleb-Ahmed A, Bekhouche SE, Hadid A (2015) Face spoofing detection from single images using active shape models with stasm and lbp. In: Proceeding of the Troisime conference internationale sur la vision artificielle CVA 2015
4. Choe KS, Sclafani AP, Litner JA, Yu G-P, Romo T (2004) The korean american woman's face: anthropometric measurements and quantitative analysis of facial aesthetics. Arch Facial Plast Surg 6(4):244–252
5. Dailey MN, Joyce C, Lyons MJ, Kamachi M, Ishi H, Gyoba J, Cottrell GW (2010) Evidence and a computational explanation of cultural differences in facial expression recognition. Emotion 10(6):874
6. Das S (2001) Filters, wrappers and a boosting-based hybrid for feature selection. In: ICML, vol 1, pp 74–81
7. Enlow DH, Moyers RE (1982) Handbook of facial growth. Saunders, Philadelphia
8. Farkas LG (1994) Anthropometry of the head and face. Raven Pr
9. Farkas LG, Katic MJ, Forrest CR (2005) International anthropometric study of facial morphology in various ethnic groups/races. J Craniofac Surg 16(4):615–646
10. Fu S-Y, Yang G-S, Kuai X-K (2012) A spiking neural network based cortex-like mechanism and application to facial expression recognition. Comput Intell Neurosci 2012:19
11. Fu S, He H, Hou Z-G (2014) Learning race from face: a survey. IEEE Trans Pattern Anal Mach Intell 36(12):2483–2509
12. Gao W, Cao B, Shan S, Chen X, Zhou D, Zhang X, Zhao D (2008) The cas-peal large-scale chinese face database and baseline evaluations. IEEE Trans Syst Man Cybern Part A Syst Hum 38(1):149–161
13. Gupta S, Castleman KR, Markey MK, Bovik AC (2010) Texas 3d face recognition database. In: IEEE southwest symposium on image analysis & interpretation (SSIAI), 2010, IEEE, pp 97–100
14. Gutta S, Wechsler H, Phillips PJ (1998) Gender and ethnic classification of face images. In: Third IEEE international conference on automatic face and gesture recognition, 1998. Proceedings. IEEE, pp 194–199
15. Gutta S, Huang JR, Jonathon P, Wechsler H (2000) Mixture of experts for classification of gender, ethnic origin, and pose of human faces. IEEE Trans Neural Netw 11(4):948–960

16. Guyomarc'H P, Dutailly B, Charton J, Santos F, Desbarats P, Coqueugniot H (2015) Anthropological facial approximation in three dimensions (afa3d): computer-assisted estimation of the facial morphology using geometric morphometrics. J Forensic Sci 59(6):1502–1516
17. He X, Yan S, Hu Y, Niyogi P, Zhang HJ (2005) Face recognition using laplacianfaces. IEEE Trans Pattern Anal Mach Intell 27(3):328–340
18. Hosoi S (2010) Face authentication apparatus, person image search system, face authentication apparatus control program, computer-readable recording medium, and method of controlling face authentication apparatus, uS Patent App 13/059,424
19. Huang Y, Yao H, Zhao S, Zhang Y (2017) Towards more efficient and flexible face image deblurring using robust salient face landmark detection. Multimed Tools Appl 76(1):123–142
20. Jianwen H, Lihua W, Longguang L, Shourong C (2010) Analysis of morphous characteristics of facial reconstruction and the five organs in chinese north five national minorities crowd. J Chongqing Med Univ 35(2):297–303
21. Keli Y, Lianbin Z, Yonglan L, Huanjiu X, Sunhua L (2016) Differences of head-face morphological traits between southern and northern chinese han. Acta Anat Sin 47(3):404–408
22. Lin H, Lu H, Zhang L (2006) A new automatic recognition system of gender, age and ethnicity. In: 2006 6th world congress on intelligent control and automation, vol 2. IEEE, pp 9988–9991
23. Lu X, Chen H, Jain AK (2006) Multimodal facial gender and ethnicity identification. In: International conference on biometrics. Springer, pp 554–561
24. Milborrow S Active shape models with stasm, Stasm Version 3
25. Ou Y, Wu X, Qian H, Xu Y (2005) A real time race classification system, IEEE
26. Peng H, Long F, Ding C (2005) Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy. IEEE Trans Pattern Anal Mach Intell 27(8):1226–1238
27. Phillips PJ, Wechsler H, Huang J, Rauss PJ (1998) The feret database and evaluation procedure for face-recognition algorithms. Image Vis Comput 16(5):295–306
28. Qiu X, Sun Z, Tan T (2006) Global texture analysis of iris images for ethnic classification. In: International conference on biometrics. Springer, pp 411–418
29. Qiu X, Sun Z, Tan T (2007) Learning appearance primitives of iris images for ethnic classification. In: 2007 IEEE international conference on image processing, vol 2. IEEE, pp II–405
30. Robinette KM, Daanen H, Paquet E (1999) The caesar project: a 3-d surface anthropometry survey. In: Second international conference on 3-D digital imaging and modeling, 1999. Proceedings. IEEE, pp 380–386
31. Sauter SL, Murphy LR, Hurrell JJ (1990) Prevention of work-related psychological disorders: a national strategy proposed by the national institute for occupational safety and health (niosh). Am Psychol 45(10):1146
32. Semwal VB, Raj M, Nandi GC (2015) Biometric gait identification based on a multilayer perceptron. Robot Auton Syst 65:65–75
33. Semwal VB, Mondal K, Nandi GC (2017) Robust and accurate feature selection for humanoid push recovery and classification: deep learning approach. Neural Comput Appl 28(3):565–574
34. Semwal VB, Singha J, Sharma PK, Chauhan A, Behera B (2017) An optimized feature selection technique based on incremental feature analysis for bio-metric gait data classification. Multimed Tools Appl 76(22):24457–24475
35. Siffert W, Forster P, Jöckel K-H, Mvere DA, Brinkmann B, Naber C, Crookes R, HEYNS ADP, EPPLEN JT, Fridey J et al (1999) Worldwide ethnic distribution of the g protein $\beta 3$ subunit 825t allele and its association with obesity in caucasian, chinese, and black african individuals. J Am Soc Nephrol 10(9):1921–1930
36. Snow CC, Gatliff BP, Mcwilliams KR (2010) Reconstruction of facial features from the skull: an evaluation of its usefulness in forensic anthropology. Am J Phys Anthropol 33(2):221–227
37. Strom MA, Zebrowitz LA, Zhang S, Bronstad PM, Lee HK (2012) Skin and bones: the contribution of skin tone and facial structure to racial prototypicality ratings. PloS One 7(7):e41193
38. Yao Y-G, Nie L, Harpending H, Fu Y-X, Yuan Z-G, Zhang Y-P (2002) Genetic relationship of chinese ethnic populations revealed by mtdna sequence diversity. Am J Phys Anthropol 118(1):63–76
39. Zhang H, Ding M, Jiao Y, Wang X, Yan Z, Jin G, Meng X, Bai C, Lu Z, Chen R (1997) A dermatoglyphic study of the chinese population dermatoglyphics cluster of fifty-two nationalities in China. Acta Genet Sin 25(5):381–391
40. Zhang H, Sun Z, Tan T, Wang J (2011) Ethnic classification based on iris images. In: Chinese conference on biometric recognition, Springer, pp 82–90
41. Zhang Y, Wang S, Phillips P, Ji G (2014) Binary pso with mutation operator for feature selection using decision tree applied to spam detection. Knowl-Based Syst 64:22–31

42. Zhao S, Yao H, Sun X (2013) Video classification and recommendation based on affective analysis of viewers. Neurocomputing 119:101–110
43. Zhao S, Yao H, Gao Y, Ji R, Ding G (2017) Continuous probability distribution prediction of image emotions via multitask shared sparse regression. IEEE Trans Multimed 19(3):632–645
44. Zhuang Z, Bradtmiller B (2005) Head-and-face anthropometric survey of us respirator users. J Occup Environ Hyg 2(11):567–576
45. Zhuang Z, Landsittel D, Benson S, Roberge R, Shaffer R (2010) Facial anthropometric differences among gender, ethnicity, and age groups. Ann Occup Hyg. meq007

Cunrui Wang received the B.S. degrees in computer science and technology from Dalian Nationalities University, Dalian, P. R. China, in 2002, and received the M.S. degrees in computer software and theory from Northeastern University, Shenyang, P. R. China, in 2005. Currently, he is a PhD candidate in the Institute of System Science, Northeastern University. His research interests include biometrics identification and data mining.

Qingling Zhang received the B.Sc. and M.Sc. degrees from the Mathematics Department and the Ph.D. degree from the Automatic Control Department of Northeastern University, China, in 1982, 1986 and 1995, respectively. Since1997, he has been a Professor with Northeastern University. He is also a member of the University Teaching Advisory Committee of National Ministry of Education. He has published six books and more than 230 papers about control theory and applications. Dr. Zhang received 14 prizes from central and local governments for his research. He has also received the Golden Scholarship from Australia in 2000. His current research interests cover fuzzy system, singular systems, data-driven control, and robust control.

Xiaodong Duan is the director of Dalian Key Lab of Digital Technology for National Culture, Dalian Nationalities University, Dalian, P. R. China. He received the the B.S. degrees in computer science and technology from Nankai University, Tianjin, P. R. China in 1985, and received the M. S. and Ph.D. degrees in applied mathematics and computer software and theory from Northeastern University, Shenyang, P. R. China in 1988 and 2001, respectively. His research interests include pattern recognition and data mining.

Jianhou Gan received the B.S. degree in computer science education and M.s. degree in mathematics from Yunnan Normal University, China, in 1998 and 2004. He received Ph.D. degree in computational metallurgy from Kunming University of Science and Technology, China, in 2016. He is the vice director of Key Laboratory of Educational Informatization for Nationalities, Yunnan Normal University, Kunming, P.R. China. His current research interests include knowledge engineering and educational informatization for nationalities.