



A novel facial attractiveness evaluation system based on face shape, facial structure features and skin

Jian Zhao¹ · Miao Zhang¹ · Chen He¹ · Xie Xie¹ · Jiaming Li¹

Received: 1 January 2020 / Revised: 5 April 2020 / Accepted: 11 April 2020 / Published online: 4 June 2020
© Springer Nature B.V. 2020

Abstract

Facial attractiveness is an important research direction of genetic psychology and cognitive psychology, and its results are significant for the study of face evolution and human evolution. However, previous studies have not put forward a comprehensive evaluation system of facial attractiveness. Traditionally, the establishment of facial attractiveness evaluation system was based on facial geometric features, without facial skin features. In this paper, combined with big data analysis, evaluation of face in real society and literature research, we found that skin also have a significant impact on facial attractiveness, because skin could reflect age, wrinkles and healthful qualities, thus affected the human perception of facial attractiveness. Therefore, we propose a comprehensive and novel facial attractiveness evaluation system based on face shape structural features, facial structure features and skin texture feature. In order to apply face shape structural features to the evaluation of facial attractiveness, the classification of face shape is the first step. Face image dataset is divided according to face shape, and then facial structure features and skin texture features that represent facial attractiveness are extracted and fused. The machine learning algorithm with the best prediction performance is selected in the face shape structural subsets to predict facial attractiveness. Experimental results show that the facial attractiveness evaluation performance can be improved by the method based on classification of face shape and multi-features fusion, the facial attractiveness scores obtained by the proposed system correlates better with human ratings. Our evaluation system can help people project their cognition of facial attractiveness into artificial agents they interact with.

Keywords Facial attractiveness · Face shape · Skin texture feature · Facial structure features · Features fusion

Introduction

An evolutionary theory assumes that perception and preferences guide biologically and socially functional behaviours (McArthur and Baron 1983). In our past

evolutionary history, power might be the basic condition that dominated individual reproduction and evolution. But with the improvement of material living standards and the richness of human spiritual activities, facial attractiveness plays an important role in many social endeavors. People tend to prefer attractive faces, so that those people have better partners and jobs, further an advantage will accrue to those in the process of evolution (Little et al. 2011; Balas et al. 2018). The study of facial attractiveness in human society has been widely researched by philosophers, psychologists, medical scientists, computer scientists, and biologists (Bashour 2006; Kagian et al. 2007; Ibáñez-Berganza et al. 2019; Jones and Jaeger 2019). Facial attractiveness is an eternal topic in human society, and the pursuit of beauty has never stopped. Even infants prefer to look at highly attractive faces (Langlois and Roggman 1990). As the demand for esthetic surgery has increased tremendously over the past few years, a comprehensive and

✉ Jian Zhao
zjctec@nwnu.edu.cn

✉ Chen He
chenhe@nwnu.edu.cn

Miao Zhang
844828152@qq.com

Xie Xie
471573313@qq.com

Jiaming Li
963509678@qq.com

¹ School of Information Science and Technology, Northwest University, Xi'an 710127, China

in-depth understanding of beauty is becoming crucial for medical settings. Recently, computer-based facial beauty analysis has become an emerging research topic, which has many potential applications, such as plastic surgery and entertainment industry (Liang and Jin 2013; Zhang et al. 2016).

The history of facial attractiveness study can be divided into several stages, from the numerical descriptions proposed in early explorations to the measurement-based criteria used in aesthetic surgery, and from the hypotheses proposed by modern psychologists to computer models for facial attractiveness evaluation (Cula et al. 2005; Atiyeh and Hayek 2008; Little et al. 2011; Gan et al. 2014). Facial attractiveness cognition includes complicated mental activities, which may change with times and individuals. However, for a specific history and cultural tradition, it was found that people used consistent criteria in their evaluation of facial attractiveness (Little et al. 2011). The findings of studies from different fields are complementary and deepen our understanding of facial beauty cognition.

Despite that, there have been numerous research activities on facial attractiveness analysis and facial attractiveness has been studied for many years, most of the traditional studies have only involved facial attractiveness in relation to geometric features and facial components, and the evaluation of facial attractiveness mostly depends on facial geometric features (DeBruine et al. 2007; Fan et al. 2012; Hong et al. 2017; Zhang et al. 2017a). Through big data analysis, evaluation of face in real society and literature research, we found that the facial skin appearance also affected the human perception of attractiveness. Since many researchers have proposed a connection between facial attractiveness and healthful qualities, the health of facial skin might be an appearance characteristic that positively influences facial attractiveness evaluation (Fink et al. 2001, 2012; Foo et al. 2017; Tan et al. 2018). At the same time, skin colour and skin texture are important factors that reflect the health of facial skin (Jones et al. 2004; Fink et al. 2006; Stephen et al. 2012).

The cross research of cognitive psychology and artificial intelligence technology has made remarkable achievements in many field. Dasdemir et al. (2017) investigated the brain networks during positive and negative emotions for different types of stimulus, and found that video content was the most effective part of a stimulus. Kumar et al. (2019) proposed a graphical based bottom up approach to point up the salient face in the crowd or in an image having multiple faces. Wang and Hu (2019) presented an automated gender recognition system based on Electroencephalography (EEG) data of resting state from twenty-eight subjects. With the advances of pattern analysis and machine learning technology, the research on facial attractiveness is no longer limited to subjective psychological cognition, facial

beauty can be predicted and improved by machine learning methods (Eisenthal 2006). While the cognition of facial beauty is easy for humans to judge, it is ambiguous for machines. Given its broad effects on human behavior and strong implications for social development, it is important to establish a comprehensive and accurate evaluation system for facial attractiveness (Ibáñez-Berganza et al. 2019). This need becomes more imperative with the large scale of face images and the increasing popularity of machine learning methods.

In this paper, we propose to establish a comprehensive and novel facial attractiveness evaluation system based on face shape structural feature, facial structure features and skin texture feature. Meanwhile, we apply the multi-features fusion method to fuse the extracted facial features. The experimental results show that the performance of our evaluation system is better than other predictors. The system proposed in this paper can extract the facial structural features and skin texture features that represent the facial attractiveness, and carry out quantitative analysis of facial attractiveness on the basis of the classification of face shape, so that the computer can objectively predict the facial attractiveness. The research achievement of this paper can provide service and reference for entertainment, social and plastic surgery. The rest of the paper is organized as follows. A brief review of literatures about facial attractiveness is presented in “Literature review” section. “Proposed facial attractiveness evaluation system based on shape, structure features and skin” section describes the proposed system in this work, and “Experimental results of the proposed system” section presents the experimental results. Finally, conclusions are drawn in “Conclusion” section.

Literature review

Facial attractiveness is affected by a diverse set of features, including geometric features, skin color, face shape and so on. They are extracted from the face images and used as features to establish predictive models for evaluating facial attractiveness. Facial components and facial geometric features are widely studied because they are not easily altered. Zhang et al. (2017a, b) recorded the binocular eye movements of participants during the judgment task, they found that the area of the nose was vital in the judgment of facial attractiveness. The ratio of the upper and lower lip, and the scleral color all affect facial attractiveness (Gründl et al. 2012; Popenko et al. 2017). Attractive female faces are consistent in many facial features, such as large eyes, slender eyebrows, small nose and so on (Baudouin and Tiberghien 2004; Schmid et al. 2008). At the same time, there is a huge difference in facial ratios and angles

between attractive and anonymous women (Milutinovic et al. 2014; Penna et al. 2015). Leite and De Martino (2015) revealed that the combination of the geometric features and skin texture modifications obtained the most significant facial beauty enhancement. Our previous research also found that different eyebrow and eye combinations have different effects on facial attractiveness (Zhao et al. 2019d). Accurate localization of facial landmarks is a fundamental step in measurement of facial geometric features. However, previous facial landmarks models did not consider the influence of forehead landmarks (Köstinger et al. 2011; Mu 2013; Chen et al. 2015; Liang et al. 2018). Therefore, in our previous study, we proposed an 81-points facial attractiveness landmarks model based on the combination of facial geometric features and illumination model (Zhao et al. 2019b).

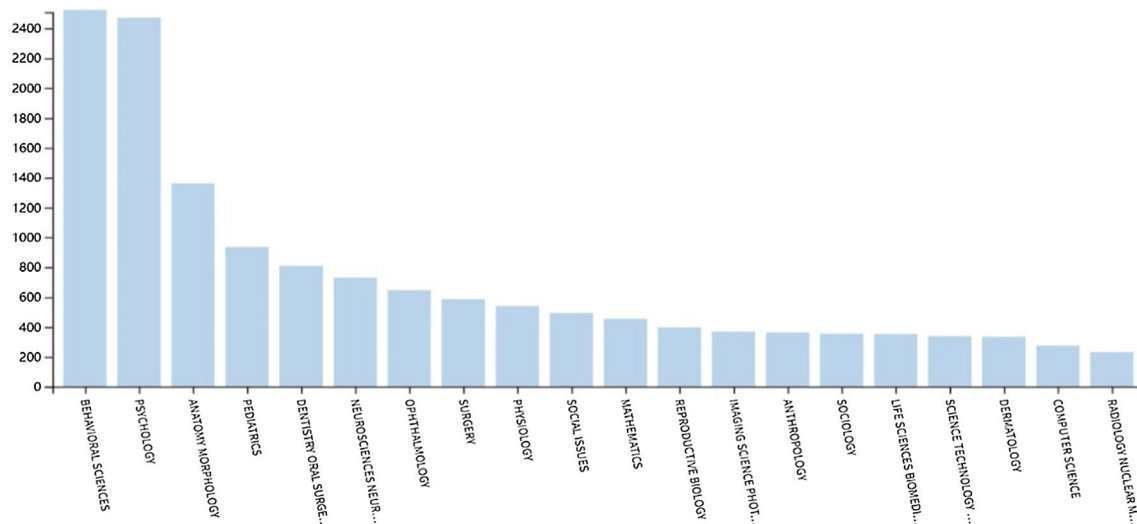
Face shape is the shape of the curve formed by the contour of the face. Considering the latest research progresses of machine learning, particularly in using deep convolutional neural networks (CNNs) in image classification (Szegedy et al. 2016), in our previous study, the pretrained Inception v3 model was used to realize face shape classification (Zhao et al. 2019a). Using CNNs in face shape classification could potentially improve classification accuracy without the need to handpick specific features extracting from the face images and then use in training the model (Zhao et al. 2019c). Meanwhile, deep-learned features may contain some important features that are not defined in the hand-crafted feature sets. Classification for face shapes can help the computer to establish recommendation system that automatically matches the appropriate hairstyles and accessories according to the face shapes (Pasupa et al. 2019). The past classification methods were mostly traditional classifiers that need to be pretrained with manually labeled geometric features (Li et al. 2013; Sunhem and Pasupa 2016). The classification accuracy of these traditional classifiers was not high enough.

By counting a large number of face image data, more than 70% of facial regions are skin-color areas, and a few of them are non-skin-color areas (Fink et al. 2001). At the same time, based on the statistics of the literatures about facial attractiveness in the web of science, we found that facial skin also has a great influence on facial attractiveness (Date: November 12, 2019). The search terms we used included ‘facial attractiveness’ and ‘facial beauty’, and 3748 results were obtained. First of all, we visualized these search results according to research direction and publication year. The research directions and publication years were only the first 20 items, as shown in Fig. 1. According to the statistical results, there was a growing literature on facial attractiveness, and the research directions involved many fields, especially behavioral science and psychology, but there was still little research in the field of computer

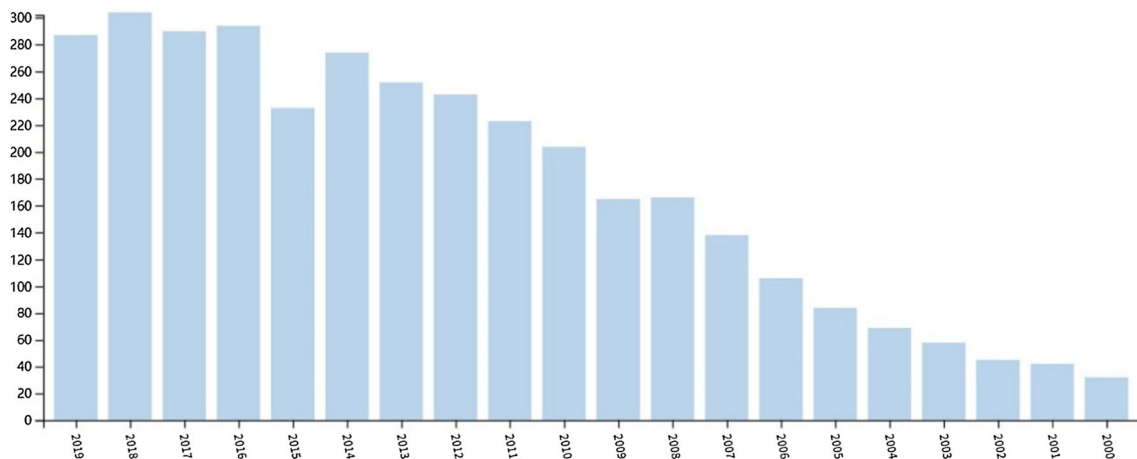
science. Secondly, we searched within results with the keyword ‘skin’, and 418 results were obtained. The retrieval results were also visualized according to research direction and publication year, as shown in Fig. 2. According to the result, there were more and more literatures about the influence of skin on the facial attractiveness, which indicated that researchers have noticed that skin has a great influence on the facial attractiveness. However, the research directions were also mostly limited to psychology and behavioral science, and there was very little research in the field of computer science. Traditionally, the evaluation of facial attractiveness depended on facial geometric features (Green et al. 2008; Fan et al. 2012; Zhang et al. 2017a). A facial beauty score can be calculated automatically by learning of the relation of ratio features and human ratings (Hong et al. 2017). Table 1 summarizes the relevant studies. In this paper, we propose to establish a comprehensive and novel facial attractiveness evaluation system based on face shape, facial structure information and skin texture.

Proposed facial attractiveness evaluation system based on shape, structure features and skin

A flowchart of the proposed facial attractiveness evaluation system is illustrated in Fig. 3. The system includes two parts: training and testing. In training, firstly, 81 facial landmarks are located in training face images, and then the classification of face shape is carried out. The facial structure features and skin texture features are extracted from each training image, and the features are fused. According to the classification of face shape, face dataset representing facial attractiveness is divided into five subsets of face shape structure, and the subsequent evaluation and prediction of facial attractiveness is limited to a small dataset. Pearson correlation coefficients are obtained by analyzing machine scoring of attractiveness measures and human ratings, and the machine learning method corresponding to the highest Pearson correlation coefficient is selected to establish facial attractiveness evaluation model for each face shape structure. In testing, classification of face shape is the most important step. Then facial attractiveness features are extracted and fused. Finally, the machine learning method corresponding to the highest Pearson correlation coefficient is selected as the facial attractiveness evaluation method in the corresponding face shape dataset. Finally, the attractiveness score for the test face is obtained.



(a) Visualize search results according to research direction



(b) Visualize search results according to publication year

Fig. 1 The search results about ‘facial attractiveness’ and ‘facial beauty’ in the web of science

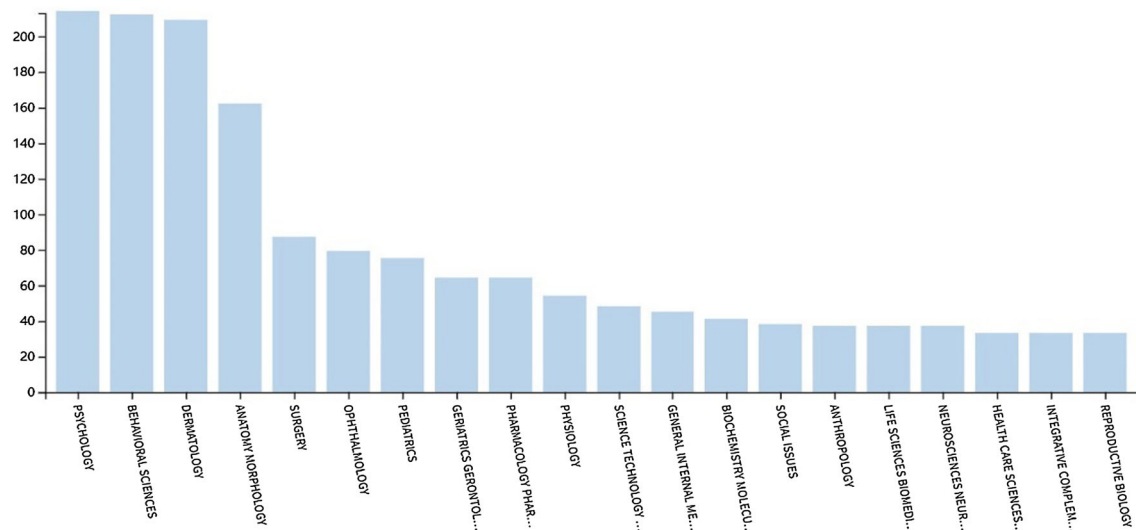
Training procedure of proposed attractiveness evaluation system

Facial landmarks

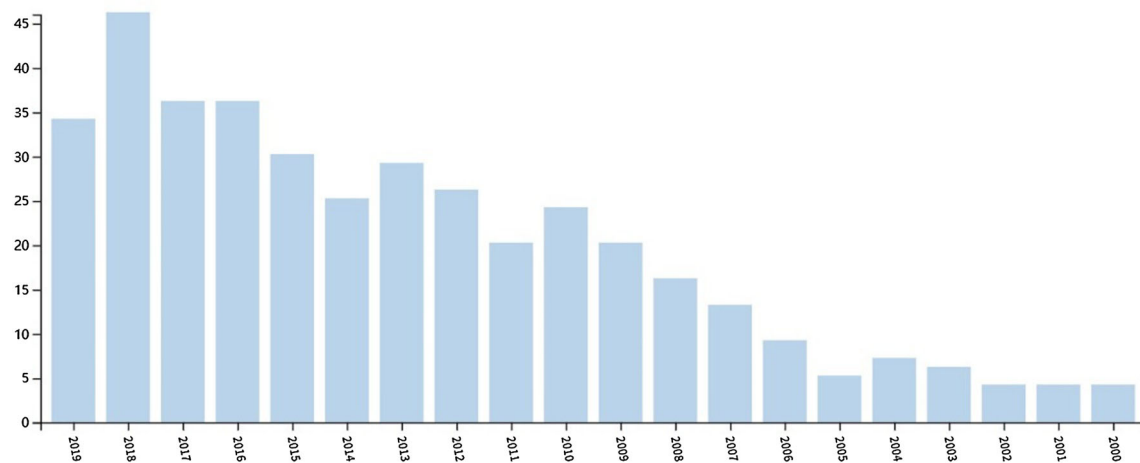
Most of the existing facial attractiveness feature landmarks templates based on computer research directly utilize successful face recognition templates (Köstinger et al. 2011; Liang and Jin 2013). However, face recognition template is a geometric template of point distribution model, the facial attractiveness landmarks template should be a geometric template combining point, line and surface. Many factors, such as the proportion of facial components, the proportion of each organ to the face area and the shape of facial components, should be considered when evaluating the facial attractiveness. These factors are not

particularly important in face recognition, but have a great influence on the evaluation of facial attractiveness.

Our previous research used transfer learning to transform the face template of face recognition into facial attractiveness template after modification according to facial attractiveness features, and built a unified 81 landmarks template of facial attractiveness (see Fig. 4) (Zhao et al. 2019b). Among them, 30 landmarks represent the face contour, 51 landmarks represent the facial components (10 landmarks represent the eyebrows, 12 landmarks represent the eyes, 9 landmarks represent the nose, 20 landmarks represent the lips, and the ears are not tested because they are covered by the hair in most face images).



(a) Visualize search results according to research direction



(b) Visualize search results according to publication year

Fig. 2 The refined results with the keyword ‘skin’ based on the search results in Fig. 1**Table 1** Summary of previous studies for evaluating facial attractiveness

Feature(s)	Measurement(s)	Researcher
Facial skin	Skin colour homogeneity	Fink et al. (2012)
Anthropometric proportions	Golden ratio, horizontal thirds, vertical fifths	Milutinovic et al. (2014)
Skin, shape	Skin colour, face shape	Carrito et al. (2016)
Landmark distance	Geometric features	Zhang et al. (2017)
Facial proportions	Four ratio feature sets	Hong et al. (2017)
Facial proportions	The proportions of face components	Przylipek et al. (2018)
Facial landmarks, face shape	Face shape, the geometric parameters of the eyebrow	Zhao et al. (2019a, b, c, d)

Face shape

In our study, we only focus on the five face shapes based on the profile features of Asian faces and previous materials (Li et al. 2013; Bansode and Sinha 2016; Sunhem and

Pasupa 2016): (1) heart-shaped; (2) oblong; (3) oval; (4) round; (5) square. The representative female and male faces for the five face shapes and their main features are shown in Table 2.

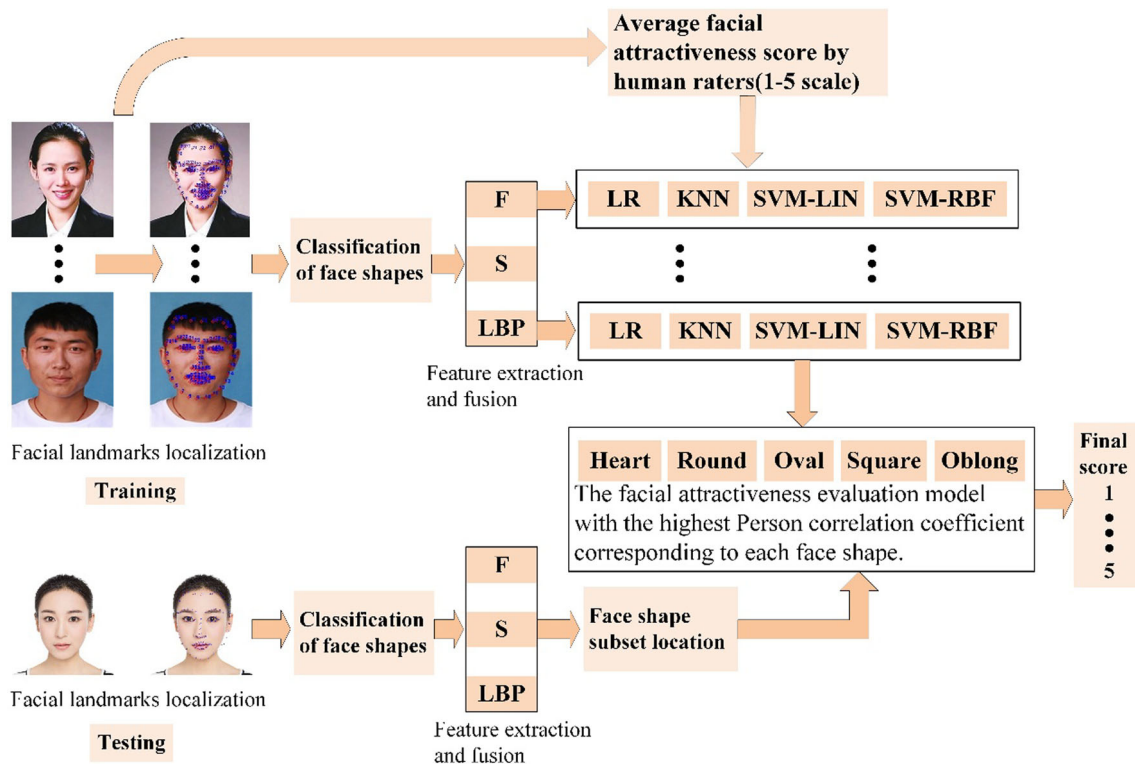


Fig. 3 Proposed system for facial attractiveness evaluation

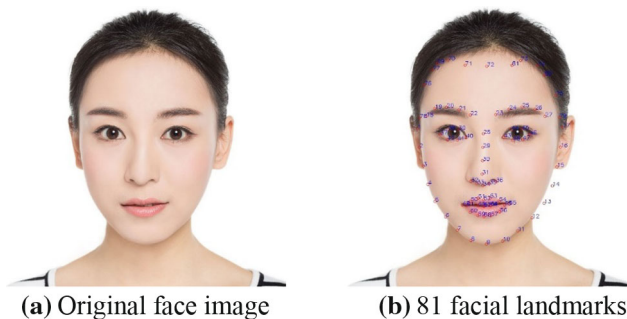


Fig. 4 Facial landmarks localization

Generally speaking, it is not easy to classify face shapes. Some face shapes features are not obvious, and different people may classify them as different face shapes when classifying them. With the advent of retrainable custom image classifiers, it has become a lot easier to make specialized image classifiers like face shape classifiers. Inception v3 is an effective retrained image classifier (Szegedy et al. 2016). By training face shape classifiers with deep learning methods, we can use some implicit characteristics which are different from previous approaches based on facial geometric features (Bansode and Sinha 2016; Pasupa et al. 2019). It can minimize human subjective bias caused by the individual differences, and makes the classification process automatic, faster and have a better consistency.

In our study, we collaborate with the College of Life Sciences, to collect face images, and get the faces attributed as one of the five shapes by human annotators, namely, heart, oblong, oval, round, and square face shapes. And the whole face database has a total of 600 images. We trained the face shape classifier using the facial features used in our previous studies (Zhao et al. 2019a). The training set sizes to train classifiers are 100, 200, 300, 400, and 500 images, respectively. For each training set size, the overall accuracy is recorded for all 500 images by using LDA, SVM-LIN, SVM-RBF, MLP, KNN classifiers, and the retrained Inception v3 model (Sunhem and Pasupa 2016). Table 3 shows the overall accuracy of different training sets by using the other classifiers and the retrained Inception v3 model. The overall accuracy of the retrained Inception v3 model ranges from 91.0% up to 93.0% while the other classifiers only range from 38.4 to 66.2% depending on the size of the training sets used. The results indicate that the retrained Inception v3 model outperform the other classifiers models.

Facial structure features

In the study of facial attractiveness, facial geometric features consider the ratio of the distance between facial landmarks, but facial components, the proportion of each organ to the face area also need to be considered (Popenko

Table 2 The representative faces for the five face shapes and their main features

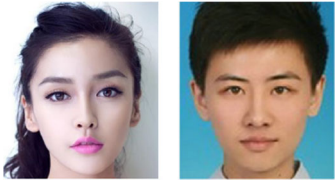




Face shape	Representative faces	Main features
Heart		Also known as the inverted triangle face, characterized by a broad forehead, narrow jaw, and a pointed chin. Heart-shaped face is one of the ideal face shapes.
Oblong		Face shape is thin and long. Forehead, cheekbone, jaw width is basically equal, but the face width is less than two-thirds of the face height. Cheeks are thin and narrow, the width of the face is significantly less than the longitudinal length, and the cheeks are sunken.
Oval		Forehead and cheekbone width are approximately equal to face width, a little wider than the jaw. The outline is curved, face width is about two-thirds of the face height.
Round		The width of forehead, cheekbone and jaw is basically the same. Rounded chin with few angles. The facial muscles are plump and lack of stereoscopic sense.
Square		Angular face with a strong jaw line, broad forehead, square and short chin. The face width is proportional to face height.

Table 3 Overall accuracy of different training size

Training size	100 (%)	200 (%)	300 (%)	400 (%)	500 (%)
MLP	38.4	46.6	48.9	50.2	52.3
SVM-RBF	45.6	48.2	49.0	49.8	51.2
SVM-LIN	47.8	49.1	53.0	54.2	55.0
LDA	57.5	58.6	60.2	64.3	65.2
KNN	43.9	51.6	55.3	57.4	66.2
Inception v3	91.0	92.1	92.1	92.8	93.0

et al. 2017; Przylipek et al. 2018). Since the facial geometric features cannot represent the whole information of a face and the human judgement on the facial attractiveness may also be disturbed by other facial features (e.g. skin),

thus constructing a facial attractiveness evaluation system directly on the geometric features is not suitable.

In our study, we propose to divide facial structure features into geometric features and triangular area features. Above all, we summarize 23 ratios and their ideal values which can reflect the facial geometric features based on previous works on facial geometric features, and they are listed in Table 4 (Schmid et al. 2008; Fan et al. 2012). The features of geometric ratio of 23 dimensions is expressed in the form of eigenvector, as shown in Eq. 1.

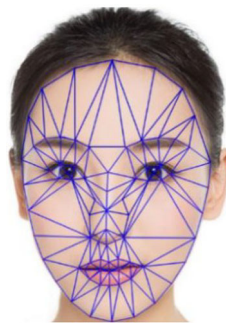
$$F = (f_1, f_2, \dots, f_n)^T \quad (1)$$

where $n = 23$.

In order to get a more accurate and complete representation of facial structure features, we need to extract not only the position and relative distance of the facial components, but also their size, length and height. The

Table 4 Summary of 23 ratios and their descriptions, ideal values

Ratio number	Description	Ideal value
<i>r</i> 1	Ear length to interocular distance	1.618
<i>r</i> 2	Ear length to nose width	1.618
<i>r</i> 3	Mideye distance to interocular distance	1.618
<i>r</i> 4	Mideye distance to nose width	1.618
<i>r</i> 5	Mouth width to interocular distance	1.618
<i>r</i> 6	Lips-chin distance to interocular distance	1.280
<i>r</i> 7	Lips-chin distance to nose width	1.280
<i>r</i> 8	Interocular distance to eye fissure width	1.618
<i>r</i> 9	Interocular distance to lip height	1.618
<i>r</i> 10	Nose width to eye fissure width	1.000
<i>r</i> 11	Nose width to lip height	1.618
<i>r</i> 12	Eye fissure width to nose-mouth distance	1.000
<i>r</i> 13	Lip height to nose-mouth distance	1.036
<i>r</i> 14	Nose width to nose-mouth distance	1.618
<i>r</i> 15	Mouth width to nose width	1.618
<i>r</i> 16	Nose length to ear length	1.000
<i>r</i> 17	Interocular distance to nose width	1.000
<i>r</i> 18	Nose-chin distance to lips-chin distance	1.618
<i>r</i> 19	Nose length to nose-chin distance	1.000
<i>r</i> 20	Length of face to width of face	1.618
<i>r</i> 21	Width of face to 5 times of nose width	1.000
<i>r</i> 22	Width of face to 5 times of eye fissure width	1.000
<i>r</i> 23	Length of face to 3 times of forehead length	1.000

Fig. 5 Schematic diagram of triangular area features

triangular area feature uses 132 triangular areas to form 132 dimensional area vectors. The triangular regions extracted by this method are labeled in Fig. 5. 132 triangular areas are constructed based on 81 facial landmarks template. These areas represent the main features of the face, such as the size of the nose, the size of a face profile, and so on. The triangular area feature of 132 dimensions is expressed in the form of eigenvector, as shown in Eq. 2.

$$S = (s_1, s_2, \dots, s_n)^T \quad (2)$$

where $n = 132$.

Skin texture features

In the study of facial attractiveness, most researchers have made relatively deep study on facial geometric features. But only focusing on the facial geometric feature will ignore the important features that affect the apparent facial attractiveness—skin texture features (Cula et al. 2005; Leite and De Martino 2015; Tan et al. 2018). Skin texture features can reflect age, wrinkles and healthful qualities, thus affecting the human perception of facial attractiveness. Because skin colour is strongly affected by the direction from which it is viewed and illuminated, we only pay close attention to skin texture features. Under the influence of external conditions, skin texture features have strong robustness. Representation and classification of skin texture has broad implications in many fields of computer vision (Singh et al. 2016). In this section, 256 dimensional feature vectors describing facial texture features are obtained based on LBP algorithm to characterize facial apparent texture features (Selim et al. 2015; Singh et al. 2016). As a simple and effective operator to extract and describe local texture features, LBP has been widely used in face recognition and other fields and achieved good results (Anbarjafari 2013; Jun et al. 2018). LBP uses Eq. 3 in actual calculation:

$$LBP(x_i, y_i) = \sum_{p=0}^{p-1} 2^p S(i_p - i_c) \quad (3)$$

where (x_i, y_i) represents the central element in a 3×3 domain, its pixel value is i_c , the values of other pixels in the neighborhood are denoted by i_p . In the basic LBP processing, p is set as 8 by default, and the symbolic function $S(x)$ is shown in Eq. 4.

$$S(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

Finally, the gray value is compared and the binary code is constructed to obtain the LBP value of each pixel. For the face images obtained in this paper, LBP texture feature vector was extracted to obtain a 256-dimensional feature vector. The facial texture feature diagram and LBP histogram are shown in Fig. 6, and the feature vector is shown in Eq. 5.

$$LBP = (z_1, z_2, \dots, z_n)^T \quad (5)$$

where $n = 256$.

Dataset and human ratings

We used the 600 face images described above as the face dataset (see “Face shape” section). All the face images are positive, neutral expression, ordinary Asian face, the age range is 20–30 years old, all the images have simple background, no accessories, reduce the interference of non-related factors, and the face images are processed by face alignment and normalization. Written informed consent was obtained from all participants. 500 of the face images were randomly selected for training and the remainder were used for testing. The dataset that we established was rated by a total of twenty participants (ten men and ten women) aged from 21 to 32 (average 24.6). None of the participants reported any known visual, tactile, olfactory, auditory, or related sensory impairment. Our ratings range from 1 to 5 are defined as follows: 1 represents very unattractive, 2 represents slightly unattractive, 3 represents mildly attractive, 4 represents slightly attractive, and 5 represents very attractive. Before scoring, we explained face shape structure rules and facial attractiveness evaluation criteria to the volunteers to determine the objective value of face shape structure and facial attractiveness quantitative evaluation. Although the evaluation ability of

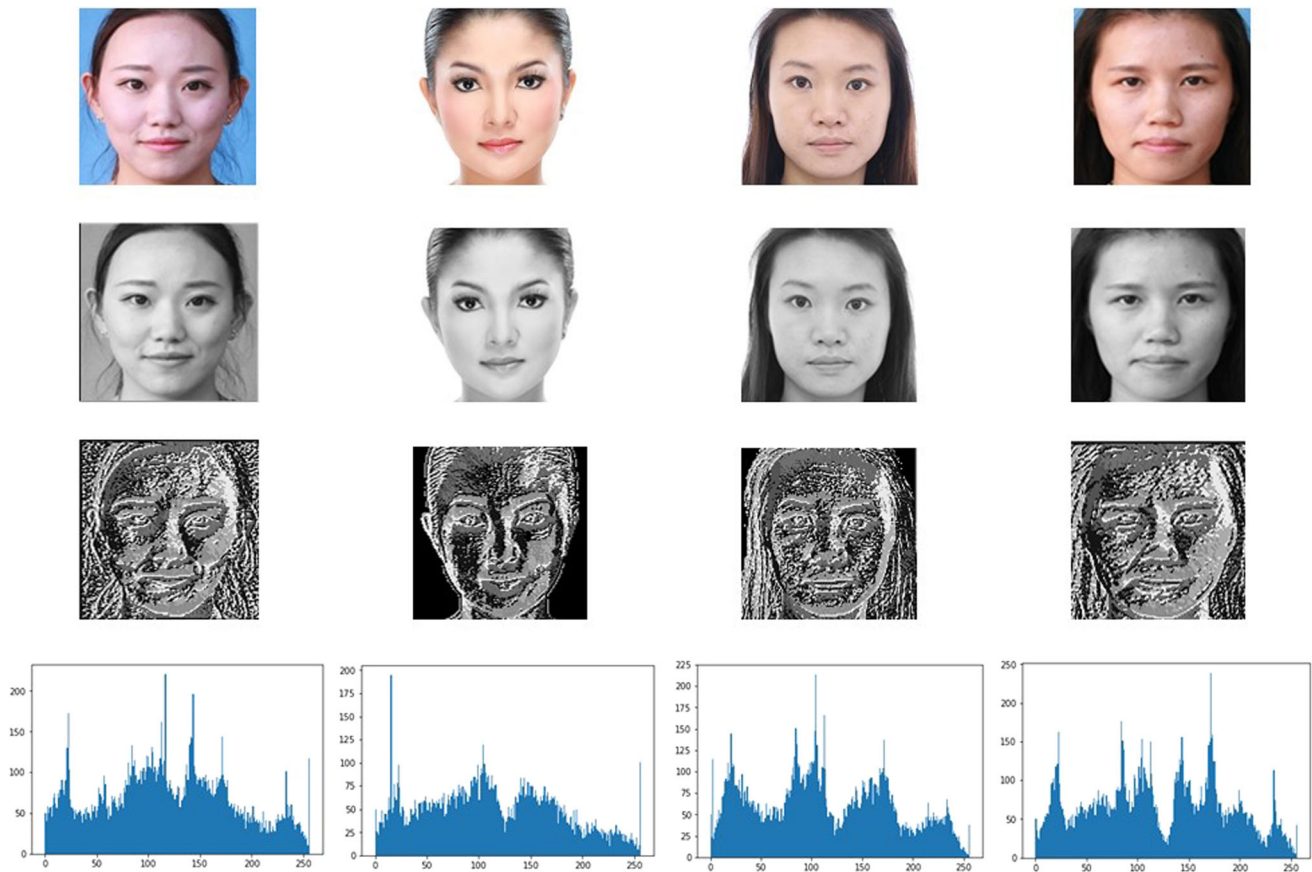


Fig. 6 LBP texture features extraction diagram

facial attractiveness of these raters has been trained and tested, there are still subjective factors, so the standardized processing of facial attractiveness evaluation data becomes indispensable (Gunes and Piccardi 2006). We used z-score normalization and linear scaling transformation to keep the rating values of each rater within the same range. Firstly, given a rating set S of a human rater, the z-score normalization formula was shown in Eq. 6.

$$z_i = \frac{s_i - \text{mean}(S)}{\text{std}(S)} \quad (6)$$

where s_i and z_i represent the i th original and normalized rating values, respectively, and $\text{mean}(S)$ and $\text{std}(S)$ represent the mean and standard deviation of a rating set S . Then, to scale a rating range from one (lower bound) to five (upper bound), the z-scores are transformed according to Eq. 7.

$$z'_i = l_b + \frac{z_i - \min(Z)}{\max(Z) - \min(Z)} \times (u_b - l_b) \quad (7)$$

where l_b and u_b represent the lower bound and upper bound of the target rating range, and $\min(\cdot)$ and $\max(\cdot)$ represent the minimum and maximum values of the given rating set, respectively. Then, the final attractiveness rating of each face image was calculated by averaging each rating from all human raters, as shown in Eq. 8.

$$Z = \text{mean}(z'_1, \dots, z'_n) \quad (8)$$

where $n = 20$, represents 20 raters.

Figure 7 shows the distribution of facial attractiveness ratings of 600 face images. The x-coordinate represents the ratings by human raters, and the y-coordinate represents the number of face images in this rating distribution range. The

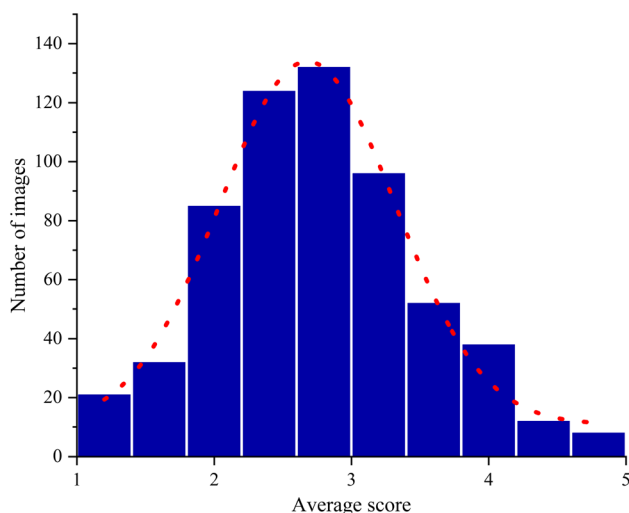


Fig. 7 Histogram and its fitting curve of average ratings by human raters

mean and standard deviation value of the ratings are 2.73 and 0.89, respectively. From the histogram and its fitting curve, it can be concluded that the ratings of face image in the dataset fit the normal distribution, and the number of face images with the ratings of 2–3 is the largest, which shows that the face dataset we collected is reasonable.

Construction of the predictor

We divided the face dataset into five face shape structural subsets, and by observing the distribution of facial attractiveness scores in the five subsets, we found that the mean ratings of heart-shaped faces (2.83) and oval faces (2.80) were higher than the mean values of the whole dataset. In addition, the proportion of faces with oval faces and heart-shaped faces in the range of 4–5 points is higher than that of other face shapes, which indicates that faces with oval faces and heart-shaped faces are more attractive than those with other face shapes. Square faces had the lowest average facial attractiveness scores (2.58), and the number of face images with lower facial attractiveness scores was significantly higher than that of faces with other face shapes, suggesting that square face shape was the least attractive face shape compared to other face shapes. According to this situation, we proposed to divide the face dataset according to the face shape to improve the performance of facial attractiveness evaluation. In other words, face dataset representing facial attractiveness was divided into five subsets of face shape according to face shape structural features in the training, and the subsequent evaluation and prediction of facial attractiveness was limited to a small dataset.

Facial geometric features show the ratio of distance between facial landmarks, triangular area features show the size of each feature of the face, and apparent texture features show the global features of the face. Considering the influence of multi-features fusion on facial attractiveness evaluation, we put the extracted geometric features, triangular area features and texture features in series, as shown in Eq. 9.

$$SSR = (F^T, S^T, LBP^T)^T \quad (9)$$

In this paper, four different machine learning methods, including linear regression (LR), linear kernel functions (SVM-LIN) and radial basis kernel function of support vector machines (SVM-RBF) and K-nearest neighbor (KNN), were used to evaluate facial attractiveness, and Pearson correlation coefficient was used as the prediction performance evaluation index. The degree of correlation between human and machine ratings of facial attractiveness can be determined by the following values: 0–0.2 indicates very weak correlation; 0.2–0.4 indicates weak correlation;

0.4–0.6 indicates moderate correlation; 0.6–0.8 indicates strong correlation; 0.8–1.0 indicates very strong correlation.

Testing procedure of proposed attractiveness evaluation system

In testing, given a new face image, classification of face shape is firstly carried out for the face image. And facial features are extracted and fused by the same procedure as in the training step. Then, in the corresponding face shape subset, several machine learning methods are used for comparative analysis with ratings by human to obtain Pearson correlation coefficient. The machine learning method corresponding to the highest Pearson correlation coefficient is selected as the facial attractiveness evaluation method to obtain the final score. In our experiment, we used 100 face images selected from our face dataset as the test set.

Experimental results of the proposed system

In the previous sections, we introduced several features that influence the evaluation of facial attractiveness. Next, in order to verify the improvement of the performance of facial attractiveness prediction by multi-features fusion, this paper evaluates the performance of facial attractiveness prediction under single feature and after mutual fusion for the three features of 23 dimensional geometric features (F), triangle area feature (S) and skin texture feature (LBP), and uses Pearson correlation coefficient as the prediction performance evaluation index. Table 5 is the experimental results of facial attractiveness evaluation performance under different facial features. Through the mutual fusion of features, we can observe the change degree of the main effect of different feature extraction, and we can find that the main effect of triangle area feature is significant, the interaction effect of triangle area feature and texture feature is significant, and when these three features are mutually fused, the Pearson correlation coefficient reaches the highest value of 0.806. Because the factors that determine facial attractiveness are diverse and complex, the effect of single facial feature is not as good as the

calculation effect after features fusion (Ibáñez-Berganza et al. 2019).

In order to improve the performance of facial attractiveness evaluation, classification of face shape is applied to facial attractiveness evaluation. After adding classification of face shape, the performance of facial attractiveness evaluation will be improved because the size of the subsequent dataset is greatly reduced. In the experiment of verifying and improving the performance of facial attractiveness evaluation, we select a total of 600 face images as dataset, and divide the dataset into five subsets by face shape structural features. The features of facial attractiveness are extracted by multi-features fusion method. LR, SVM-LIN, SVM-RBF and KNN are used to select the most relevant facial attractiveness evaluation method for different face shape. Pearson correlation coefficient is used to compare the prediction performance. As shown in Table 6, each face shape subset has relatively high correlation with the ratings by the human raters in the test set. From Table 6, we find different face shapes use different machine learning algorithms have great differences in facial attractiveness evaluation performance, which further confirms that face shape plays an important role in facial attractiveness evaluation. Different machine learning algorithms are selected for the five face shape subsets, KNN algorithm is selected for the oval face, SVM-LIN algorithm is selected for the oblong face and the heart face, SVM-RBF algorithm is selected for the round face and the square face. The machine learning algorithms with the highest Pearson correlation coefficient for the attractiveness evaluation of various face shapes are combined and

Table 6 Predictive performance of facial attractiveness for five face shape subsets

Face shape	Round	Oval	Oblong	Square	Heart
LR	0.802	0.797	0.825	0.783	0.818
KNN	0.793	0.845	0.802	0.812	0.786
SVM-LIN	0.767	0.773	0.864	0.790	0.858
SVM-RBF	0.855	0.815	0.782	0.852	0.792

The bold values are to emphasize the superiority of the results obtained and highlight the method of use

Table 5 Facial attractiveness predictive performance under different features

Predictive performance	F	S	LBP	F × S	F × LBP	S × LBP	F × S × LBP
LR	0.502	0.616	0.658	0.683	0.654	0.637	0.722
KNN	0.619	0.672	0.694	0.753	0.771	0.782	0.794
SVM-LIN	0.649	0.738	0.712	0.768	0.732	0.724	0.797
SVM-RBF	0.702	0.713	0.741	0.763	0.754	0.781	0.806

The bold values are to emphasize the superiority of the results obtained and highlight the method of use

Table 7 Comparison of facial attractiveness prediction performance before and after classification of face shape

Performance index	Pearson	MAE	RMSE
The average performance after face shape classification	0.862	0.161	0.213
The average performance before face shape classification	0.802	0.189	0.254

spliced to obtain the highest evaluation performance of facial attractiveness.

By selecting 100 face images in the face dataset for testing, the method based on classification of face shape is used to improve the evaluation performance of facial attractiveness, and the comparison results are shown in Table 7. Here, we use the method of multi-features fusion to extract facial features. According to the experimental results, Pearson correlation coefficient of facial attractiveness after classification of face shape is up to 0.862, which is improved by 7.5% compared with that before classification of face shape. In addition, MAE and RMSE also decrease, MAE = 0.161 and RMSE = 0.213, indicating that the performance of facial attractiveness evaluation is improved to a certain extent (MAE is mean absolute error, it represents the average of the absolute values of the deviation between the predicted values and human ratings in facial attractiveness evaluation, when MAE is smaller, it indicates that the prediction performance of facial attractiveness evaluation is better; RMSE is root mean squared error, it is the square root of the ratio of the sum of the squares of the difference between the predicted value and the real value to the prediction times. In facial attractiveness prediction, it refers to the deviation range between the machine ratings and the human ratings. When the RMSE value is smaller, it shows that the predicted result of facial attractiveness prediction model has better fitting accuracy with the human ratings). Meanwhile, we find that the facial attractiveness evaluation performance based on the combination of classification of face shape and multi-features fusion is better than that in the former two cases, which indicates that the system we established is reasonable and accurate.

Conclusions

In this study, we proposed a comprehensive and novel facial attractiveness evaluation system based on face shape structural features, facial structure features and skin texture feature. Classification of face shape was applied to facial attractiveness evaluation, and a facial attractiveness evaluation system based on face shape subset was constructed. At the same time, we extracted facial geometric features, triangular area features and skin texture features as features representing facial attractiveness and conducted features fusion. Finally, we selected the machine learning algorithm

with the highest Pearson correlation coefficient to predict the attractiveness. The results indicated that the machine evaluation model of facial attractiveness based on classification of face shape and multi-features fusion could improve the performance of facial attractiveness evaluation. Our facial attractiveness evaluation system can be used in many areas, such as facial beauty ranking, plastic surgery and entertainment industry. Our further works consist of determining quantitative proportions of factors that affect facial attractiveness, taking into account skin colour, and increasing the size of the face dataset.

Acknowledgements This work was supported by National Natural Science Foundation of China (Grant No. 61701401) and Natural Science Foundation of Shaanxi Province (Grant No. 2019KJXX-061).

Compliance with ethical standards

Conflict of interest The authors declare no conflict of interest.

References

- Anbarjafari G (2013) Face recognition using color local binary pattern from mutually independent color channels. *EURASIP J Image Video Process* 2013:6. <https://doi.org/10.1186/1687-5281-2013-6>
- Atiyeh BS, Hayek SN (2008) Numeric expression of aesthetics and beauty. *Aesthet Plast Surg* 32:209–216. <https://doi.org/10.1007/s00266-007-9074-x>
- Balas B, Tupa L, Pacella J (2018) Measuring social variables in real and artificial faces. *Comput Human Behav* 88:236–243. <https://doi.org/10.1016/j.chb.2018.07.013>
- Bansode NK, Sinha PK (2016) Face shape classification based on region similarity, correlation and fractal dimensions. *Int J Comput Sci Issues* 13:24–31. <https://doi.org/10.20943/ijcsi-201602-2431>
- Bashour M (2006) An objective system for measuring facial attractiveness. *Plast Reconstr Surg* 118:757–774. <https://doi.org/10.1097/01.prs.0000207382.60636.1c>
- Baudouin JY, Tiberghien G (2004) Symmetry, averageness, and feature size in the facial attractiveness of women. *Acta Psychol (Amst)* 117:313–332. <https://doi.org/10.1016/j.actpsy.2004.07.002>
- Carrito M, Santos IM, Lefevre CE et al (2016) The role of sexually dimorphic skin colour and shape in attractiveness of male faces. *Evol Hum Behav* 37:125–133. <https://doi.org/10.1016/j.evolhumbehav.2015.09.006>
- Chen F, Xu Y, Zhang D, Chen K (2015) 2D facial landmark model design by combining key points and inserted points. *Expert Syst Appl* 42:7858–7868. <https://doi.org/10.1016/j.eswa.2015.06.015>
- Cula OG, Dan KJ, Murphy FP, Rao BK (2005) Skin texture modeling. *Int J Comput Vis* 62:97–119. <https://doi.org/10.1023/B:VISI.0000046591.79973.6f>

- Dasdemir Y, Yildirim E, Yildirim S (2017) Analysis of functional brain connections for positive–negative emotions using phase locking value. *Cogn Neurodyn* 11:487–500. <https://doi.org/10.1007/s11571-017-9447-z>
- DeBruine LM, Jones BC, Unger L et al (2007) Dissociating averageness and attractiveness: attractive faces are not always average. *J Exp Psychol Hum Percept Perform* 33:1420–1430. <https://doi.org/10.1037/0096-1523.33.6.1420>
- Eisenthal Y (2006) Facial attractiveness: beauty and the machine. *Neural Comput* 18:119–142
- Fan J, Chau KP, Wan X et al (2012) Prediction of facial attractiveness from facial proportions. *Pattern Recognit* 45:2326–2334. <https://doi.org/10.1016/j.patcog.2011.11.024>
- Fink B, Grammer K, Thornhill R (2001) Human (homo sapiens) facial attractiveness in relation to skin texture and color. *J Comp Psychol* 115:92–99. <https://doi.org/10.1037/0735-7036.115.1.92>
- Fink B, Grammer K, Matts PJ (2006) Visible skin color distribution plays a role in the perception of age, attractiveness, and health in female faces. *Evol Hum Behav*. <https://doi.org/10.1016/j.evolhumbehav.2006.08.007>
- Fink B, Matts PJ, D’Emiliano D et al (2012) Colour homogeneity and visual perception of age, health and attractiveness of male facial skin. *J Eur Acad Dermatol Venereol* 26:1486–1492. <https://doi.org/10.1111/j.1468-3083.2011.04316.x>
- Foo YZ, Simmons LW, Rhodes G (2017) Predictors of facial attractiveness and health in humans. *Sci Rep*. <https://doi.org/10.1038/srep39731>
- Gan J, Li L, Zhai Y, Liu Y (2014) Deep self-taught learning for facial beauty prediction. *Neurocomputing* 144:295–303. <https://doi.org/10.1016/j.neucom.2014.05.028>
- Green RD, MacDorman KF, Ho CC, Vasudevan S (2008) Sensitivity to the proportions of faces that vary in human likeness. *Comput Human Behav* 24:2456–2474. <https://doi.org/10.1016/j.chb.2008.02.019>
- Gründl M, Knoll S, Eisenmann-Klein M, Prantl L (2012) The blue-eyes stereotype: do eye color, pupil diameter, and scleral color affect attractiveness? *Aesthet Plast Surg* 36:234–240. <https://doi.org/10.1007/s00266-011-9793-x>
- Gunes H, Piccardi M (2006) Assessing facial beauty through proportion analysis by image processing and supervised learning. *Int J Hum Comput Stud*. <https://doi.org/10.1016/j.ijhcs.2006.07.004>
- Hong YJ, Nam GP, Choi H et al (2017) A novel framework for assessing facial attractiveness based on facial proportions. *Symmetry (Basel)* 9:1–12. <https://doi.org/10.3390/sym9120294>
- Ibáñez-Berganza M, Amico A, Loreto V (2019) Subjectivity and complexity of facial attractiveness. *Sci Rep*. <https://doi.org/10.1038/s41598-019-44655-9>
- Jones AL, Jaeger B (2019) Biological bases of beauty revisited: the effect of symmetry, averageness, and sexual dimorphism on female facial attractiveness. *Symmetry (Basel)* 11:279. <https://doi.org/10.3390/SYM11020279>
- Jones BC, Little AC, Burt DM, Perrett DI (2004) When facial attractiveness is only skin deep. *Perception* 33:569–576. <https://doi.org/10.1068/p3463>
- Jun Z, Jizhao H, Zhenglan T, Feng W (2018) Face detection based on LBP. In: *ICEMI 2017—Proceedings of IEEE 13th international conference on electronic measurement and instruments*
- Kagian A, Dror G, Leyvand T et al (2007) A humanlike predictor of facial attractiveness. *Adv Neural Inf Process Syst*. <https://doi.org/10.7551/mitpress/7503.003.0086>
- Köstinger M, Wohlhart P, Roth PM, Bischof H (2011) Annotated facial landmarks in the wild: a large-scale, real-world database for facial landmark localization. In: *Proceedings of the IEEE international conference on computer vision*
- Kumar RK, Garain J, Kisku DR, Sanyal G (2019) Guiding attention of faces through graph based visual saliency (GBVS). *Cogn Neurodyn* 13:125–149. <https://doi.org/10.1007/s11571-018-9515-z>
- Langlois JH, Roggman LA (1990) Attractive faces are only average. *Psychol Sci* 1:115–121
- Leite TS, De Martino JM (2015) Improving the attractiveness of faces in images. In: *Brazilian symposium of computer graphic and image processing*
- Li L, So J, Shin H, Han Y (2013) An AAM-based face shape classification method used for facial expression recognition. *Int J Res Eng Technol* 2:164–168
- Liang L, Jin L (2013) Facial skin beautification using region-aware mask. In: *Proceedings—2013 IEEE International conference on systems, man, and cybernetics SMC 2013* (pp. 2922–2926). <https://doi.org/10.1109/SMC.2013.498>
- Liang L, Lin L, Jin L, et al. (2018) SCUT-FBP5500: a diverse benchmark dataset for multi-paradigm facial beauty prediction. In: *Proceedings—international conference on pattern recognition*
- Little AC, Jones BC, DeBruine LM (2011) Facial attractiveness: evolutionary based research. *Philos Trans R Soc B Biol Sci* 366:1638–1659. <https://doi.org/10.1098/rstb.2010.0404>
- McArthur LZ, Baron RM (1983) Toward an ecological theory of social perception. *Psychol Rev*. <https://doi.org/10.1037/0033-295X.90.3.215>
- Milutinovic J, Zelic K, Nedeljkovic N (2014) Evaluation of facial beauty using anthropometric proportions. *Sci World J*. <https://doi.org/10.1155/2014/428250>
- Mu Y (2013) Computational facial attractiveness prediction by aesthetics-aware features. *Neurocomputing*. <https://doi.org/10.1016/j.neucom.2012.06.020>
- Pasupa K, Sunhem W, Loo CK (2019) A hybrid approach to building face shape classifier for hairstyle recommender system. *Expert Syst Appl* 120:14–32. <https://doi.org/10.1016/j.eswa.2018.11.011>
- Penna V, Fricke A, Iblher N et al (2015) The attractive lip: a photomorphometric analysis. *J Plast Reconstr Aesthetic Surg*. <https://doi.org/10.1016/j.bjps.2015.03.013>
- Popenko NA, Tripathi PB, Devic Z et al (2017) A quantitative approach to determining the ideal female lip aesthetic and its effect on facial attractiveness. *JAMA Facial Plast Surg* 19:261–267. <https://doi.org/10.1001/jamafacial.2016.2049>
- Przylipek M, Przylipek J, Terlikowski R et al (2018) Impact of face proportions on face attractiveness. *J Cosmet Dermatol* 17:954–959. <https://doi.org/10.1111/jocd.12783>
- Schmid K, Marx D, Samal A (2008) Computation of a face attractiveness index based on neoclassical canons, symmetry, and golden ratios. *Pattern Recognit* 41:2710–2717. <https://doi.org/10.1016/j.patcog.2007.11.022>
- Selim M, Raheja S, Stricker D (2015) Real-time human age estimation based on facial images using uniform local binary patterns. In: *VISAPP 2015—10th international conference on computer vision theory and applications; VISIGRAPP, proceedings*
- Singh R, Shah P, Bagade J (2016) Skin texture analysis using machine learning. In: *Conference on advances in signal processing, CASP 2016*
- Stephen ID, Scott IML, Coetzee V et al (2012) Cross-cultural effects of color, but not morphological masculinity, on perceived attractiveness of men’s faces. *Evol Hum Behav*. <https://doi.org/10.1016/j.evolhumbehav.2011.10.003>
- Sunhem W, Pasupa K (2016) An approach to face shape classification for hairstyle recommendation. In: *Proc 8th international conference on advanced computational intelligence ICACI 2016* (pp. 390–394). <https://doi.org/10.1109/ICACI.2016.7449857>

- Szegedy C, Vanhoucke V, Ioffe S, et al (2016) Rethinking the inception architecture for computer vision. In: Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 2818–2826). <https://doi.org/10.1109/CVPR.2016.308>
- Tan KW, Tiddeman B, Stephen ID (2018) Skin texture and colour predict perceived health in Asian faces. *Evol Hum Behav* 39:320–335. <https://doi.org/10.1016/j.evolhumbehav.2018.02.003>
- Wang P, Hu J (2019) A hybrid model for EEG-based gender recognition. *Cogn Neurodyn* 13:541–554. <https://doi.org/10.1007/s11571-019-09543-y>
- Zhang D, Chen F, Xu Y (2016) Computer models for facial beauty analysis. Springer, Switzerland
- Zhang L, Zhang D, Sun MM, Chen FM (2017a) Facial beauty analysis based on geometric feature: toward attractiveness assessment application. *Expert Syst Appl* 82:252–265. <https://doi.org/10.1016/j.eswa.2017.04.021>
- Zhang Y, Wang X, Wang J et al (2017b) Patterns of eye movements when observers judge female facial attractiveness. *Front Psychol* 8:1–7. <https://doi.org/10.3389/fpsyg.2017.01909>
- Zhao J, Cao M, Xie X et al (2019a) Data-driven facial attractiveness of Chinese male with epoch characteristics. *IEEE Access* 7:10956–10966. <https://doi.org/10.1109/ACCESS.2019.2892137>
- Zhao J, Deng F, Jia J et al (2019b) A new face feature point matrix based on geometric features and illumination models for facial attraction analysis. *Discret Contin Dyn Syst S* 12:1065–1072. <https://doi.org/10.3934/dcdss.2019073>
- Zhao J, Xie X, Wang L et al (2019c) Generating photographic faces from the sketch guided by attribute using GAN. *IEEE Access*. <https://doi.org/10.1109/ACCESS.2019.2899466>
- Zhao J, Zhang M, He C, Zuo K (2019d) Data-driven research on the matching degree of eyes, eyebrows and face shapes. *Front Psychol* 10:1–11. <https://doi.org/10.3389/fpsyg.2019.01466>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.