Proceedings of the 2009 IEEE International Conference on Systems, Man, and Cybernetics
San Antonio, TX, USA - October 2009

Automatic Classification of Chinese Female Facial Beauty using Support Vector Machine

Huiyun Mao, Lianwen Jin*, Minghui Du
School of Electronic and Information engineering
South China University of Technology
Guangzhou, P.R. China
Huiyun.mao@gmail.com, *lianwen.jin@gmail.com, ecmhdu@scut.edu.cn

***Abstract*—Beauty is a universal concept which has long been explored by philosophers, artists and psychologists, but there are few implementations of automated facial beauty assessment in computational science. In this paper, we develop an automated Chinese female facial beauty classification system through the application of machine learning algorithm of SVM (Support Vector Machine). We present a simple but effective feature extraction for facial beauty classification. 17 geometric features are designed to abstractly represent each facial image. The experiment is based on 510 facial images, high accuracy of 95.3% is obtained for 2-level classification (beautiful or not), but the accuracy of 4-level classification is 77.9% by SVM. The results clearly show that the notion of beauty perceived by human can also be learned by machine through the employment of supervised learning techniques. Furthermore, the finding of big gap between the accuracy of 2-level classification and 4-level classification is interesting and surprising: the high accuracy naturally leads to the conclusion that there indeed exists simplicity and objectiveness underlying the judgment of aesthetical ideal facial attractiveness; In contrast, the relatively low accuracy for 4-level classification indicates that the presented simple feature vectors is not sufficient for the classification of other levels of facial attractiveness .**

Keywords—Facial beauty classification, Chinese female facial beauty, Support Vector Machine

I. INTRODUCTION

Beauty is a universal part of human experience. “There are few more pleasurable sights than a beautiful face”, “Attractive faces activate reward centers in the brain” [1, 2]. The power of beauty in human affairs has been expounded since the advent of writing, Plato understood that to be beautiful was 1 of the 3 wishes of every man, the other 2 being good health and riches acquired by honest means. Aristotle conceded beauty is a greater recommendation than any letter of introduction [3, 4]. But what is the essence of beauty? Is the judgment of beauty subjective or objective? How to define beauty? Philosophers, scientists, even ordinary people have long puzzled over these problems, and the pursing of these answers has led to a large body of ongoing research by scientists in the biological, cognitive and computational science. From evolutionary perspective, "Darwinian" ’s reasonable working hypothesis is that the psychological mechanisms underlying attractiveness judgments are adaptations that have evolved in the service of choosing a mate so as to increase gene propagation throughout evolutionary history, attractive faces advertise a "health certificate", indicating a person's "value" as a mate [5, 6], these studies focus on three major lines ---facial symmetry, averageness and secondary sex characteristics as hormone markers. And the study results show that symmetry, secondary sex characteristics have positive correlation with beauty, health as well as viable offspring, but the relation between averageness and beauty has been widely debated; In the social sciences, a long-held view is that standards of beauty are arbitrary cultural conventions [7, 3]. Even Darwin favored this view after observing large cultural differences in beautification practices [8]. However, the recent observations demonstrate that people in different cultures generally agree on which faces are attractive [9, 10, 11].

In computational science, software of constructing beautiful face emerges only a decade ago, and there are few researches by using the methods of image processing and machine learning to produce automatic human-like facial beauty assessment systems, which from other perspectives help exploring the questions such as “what is the constituents of beauty? “Is there any objective criteria underlying the perception of beauty?” “Can beauty be described with mathematical rules?” Parham et al. [12] developed an automatic facial beauty scoring system based on 8-element facial feature vectors, they use a variant of K-nearest neighbor algorithm to learn a beauty assignment function from 40 facial images rated by humans, the most high accuracy of beauty classification is 91%; H. Irem Turkmen et al. [13] developed a method that measures facial beauty based on 150 female facial images, the success ratios with KPCA and PCA are 89% and 83%, respectively; Y. Eisenthal et al. [14] collected 92 facial images for training and another independent set of 92 facial images for testing, they use standard KNN and SVM for classifier, and the accuracy of KNN and SVM is 86% and 84% respectively. However, the data sets used in these approaches are not large enough or not diversified (i.e. beauty extents in the used facial images are not dispersed enough) or both, and the set of human raters involved in some methods may not be sufficient. The improvement in such issues is important for classifier to produce more meaningful results.

978-1-4244-2794-9/09/$25.00 ©2009 IEEE

In this paper, we presents an automatic facial beauty classification method based on the geometric features related to beauty, 17-element feature vector is extracted for each given facial image, and two typical machine learning methods, SVM and C4.5, are employed for automatic beauty learning and classification. Considering system's performance and robustness, we collect a large quantity of facial images as our data sets, which are composed of 510 front-on portraits of Chinese females, with beauty extents widely dispersed, and all of the images are scored on a 4-level scale (3 most beautiful, 0 least beautiful) beforehand by human raters (the median score was used as the final measure in the training phase.) We randomly select 75% of the images as the training data set (381 images) for training SVM and C4.5 classifiers, and the remaining 25% as testing data set (127 images). The experimental results are very interesting and encouraging: the predictor's accuracy of score 3 (most beautiful level) is surprisingly high (take SVM for example, 96.8% is obtained), instead, the accuracy of other scores is relatively low. In addition, SVM outperforms C4.5. Overall, our system can provide the following advantages compared with the previous approaches: (1) The size of data sets we used is significantly larger. (2) The diversity of beauty extents in data sets is much wider. (3) The human referees are sufficient; (4) The feature and feature normalization we present is simple but effective, thus leads to high accuracy for the classification of aesthetical ideal beauty. With the advantages (1)(2)(3), our experimental results can make more sense and are more closer to real world.

The remainder of this paper is organized as follows. Section 2 describes the experimental data, the feature extraction method as well as the classification algorithms employed in the system. Section 3 gives the experimental results. And finally, Section 4 consists of a discussion of the work we presented and conclusions.

II. Facial Beauty Classification Using Geometric Features and Machine Learning Algorithm

A. Data preparation

The database we used in the experiment consists of two parts, CAS-PEAL Face dataset [15] and other facial images we collected from the Internet. The CAS-PEAL dataset is a large scale Chinese face database, which contains 99,450 samples of 1040 persons (595 males and 445 females). For the purpose of reducing the effects of facial expression and other irrelevant factors, we adopted 371 front-view female facial images with neural expression and nearly identical lighting condition from CAS-PEAL, and there are no accessories in these images. Although the number of facial images selected is large enough, however, we found the dataset lacks obvious beautiful-looking faces and poor-looking faces. Therefore, we collected other 139 female facial images from the Internet (most of them are TV/movie stars and have been generally acknowledged as beautiful females in China, and the remaining are poor-looking female faces). In total, there are 510 Chinese female facial images used as our experimental data.

Human raters, most in their twenties, were required to score each of the 510 facial images according to a scale of n=4 grades (levels) described as belows,

- 0: Poor looking,
- 1: Common looking,
- 2: Good looking,
- 3: Beautiful looking.

Fig. 1 illustrates some examples of the facial images and their respective beauty scores. In the process of human rating, we found that it is easy to achieve agreement on the scoring of beautiful looking class, but hard on the other three classes. For example, different individuals may give different scores for Fig. 1(b) and Fig. 1(c). Some people labeled the face shown in Fig. 1(c) into "good looking" class while the others scored Fig. 1(b) into "common-looking" class. The high consistency in level 3 by human rating is interesting, thus, to make the underlying regularity more clear, we design another scoring scheme thereafter, in which we simply grade the 510 samples into two classes, beautiful looking (118 faces in total) and common looking (392 faces in total) to have further investigation.

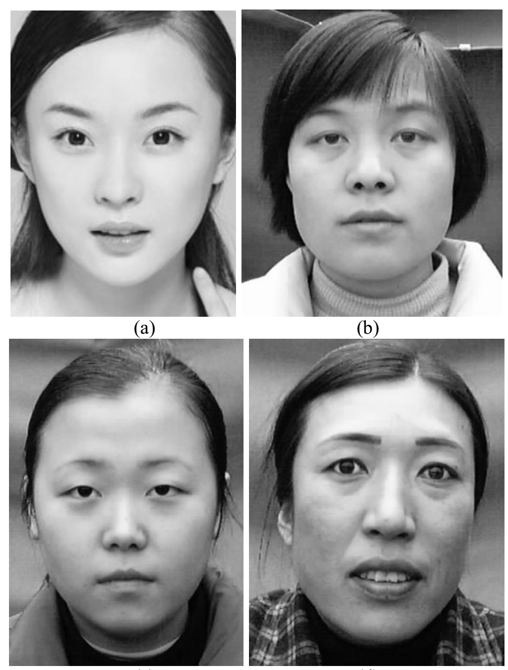

Figure 1. Some examples with different levels of beauty. (a). Beautiful looking (b) Good looking (c) Common looking (d) Poor looking

B. Feature extraction for beauty classification

Most artist and psychologists found that the contour and harmonious features in faces have great contribution to facial beauty [16, 17]. Inspired from this and through large investigation of our Chinese female beauty dataset, we extract 17 features to abstractly represent each face in our data sets. The detailed feature extraction process is described as below:

Step1. Draw 9 horizontal lines manually from top to bottom of the face through the middle point between eyebrows, upper eyelids, pupils, lower eyelids, cheekbones, base of nose, middle of lips, middle of chin, bottom of chin, respectively.

Step2. Draw 9 vertical lines from left to right of the face through outer edge of right cheekbone, outer edge of the right eye, right pupil, inner edge of right eye, middle point between two eyebrows, inner edge of left eye, left pupil, outer edge of left eye, outer edge of left cheekbone, respectively.

Step3. Locate 4 points inside face: outer edge of left nostril, right edge of left nostril, right edge of mouth, left corner of mouth.

Step4. Locate 12 points on the edges of face: right and left edge point of face at temple level, right and left edge points at nostril level, right and left edge points at mouth level, right of chin, left of chin, the remaining 4 points can be randomly selected on the bottom of jaw.

Step5. Calculate the distances between the points above labeled, as shown in Fig. 2(b).

Step6. Normalize the feature distances by the distance between inner edges of eyes.

Fig. 2(a) shows the 16 feature points after the process of step 1 ~ step4. Fig. 2(b) illustrates the 17 features extracted based on the feature points, which we denoted as F1, F2, …, F17 respectively. The physical meaning of each feature is summarized in Table Ⅰ.

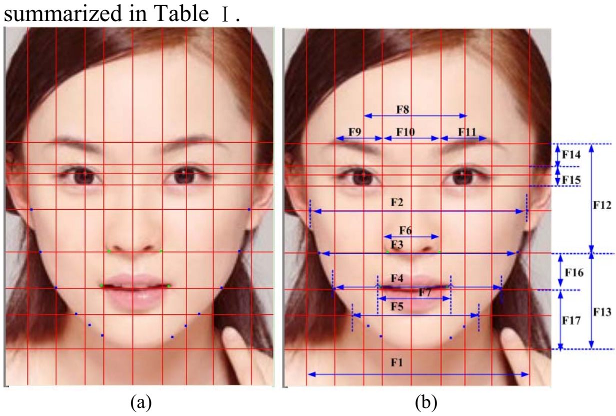

Figure 2. Feature extraction for beauty classification. (a). The 16 feature points after the process of horizontal and vertical dividing of face image; (b). The 17 features we extracted for beauty learning and classification

It is worthwhile to mention that some conventional features related to the concept of beauty, such as the Golden Ratio [16], can be easily calculated from our 17 features shown in Table Ⅰ. For example, the ratio of distance between eye and mouth to distance between mouth and chin can be calculated by (F12+F16)/F17, the ratio of distance between eyes and nose to distance between nose and mouth can be calculated by (F12-F14-F15/2)/F16, the ratio of distance between mouth and chin to the distance between nose and mouth can be calculated by F17/F16, and so on. From general aesthetical viewpoint [16], the closer these proportions are to the golden ratio (0.618), the more beautiful a face would be.

TABLE Ⅰ. 17 GEOMETRIC FEATURES EXTRACTED FOR THE CLASSIFICATION OF FACIAL BEAUTY

Feature	Description
F1	Horizontal face length at temple level
F2	Horizontal face length at cheekbone level(ears are excluded)
F3	Horizontal face length at cheekbone level
F4	Horizontal face length at mouth level
F5	Horizontal length of chin
F6	Horizontal length of nose
F7	Horizontal length of mouth
F8	Horizontal distance between pupils
F9	Horizontal length of left eye
F10	Horizontal distance between inner edges of eyes
F11	Horizontal length of right eye
F12	Vertical distance from middle point between eyebrows to the bottom of nose
F13	Vertical distance from nose bottom to face bottom
F14	Vertical distance between eyebrows and eyes
F15	Vertical height of eyes
F16	Vertical distance between nose and mouth
F17	Vertical distance between mouth and chin

C. *Beauty Classification using Supervised Machine Learning Algorithm*

From a new perspective other than psychological and social science, we regard the assessment of facial beauty as a supervised machine learning problem. Therefore we formulate the facial beauty evaluation as a classification problem where classes correspond to beauty scores (class 0: poor looking, class 1: Common looking, class 2: Good looking, class 3: beautiful looking). Each face instance is represented as a 17-element feature vector extracted from section II.

Support Vector Machine (SVM) [19] is a well known machine learning algorithm for its generalization performance and ability in handling non-linear classification problem. Consider a two class classification problem. Let $\{(\mathbf{x}_1, y_1), (\mathbf{x}_2, y_2), \cdots (\mathbf{x}_N, y_N)\}$ be a two-class training dataset, with $\mathbf{x}_i$ a training feature vector and their class labels $y_i \in \{-1,1\}$, and N is the total number of training samples. The SVM attempts to find an optimal separating hyperplane to maximally separate two classes of training samples, and learn a decision function. The corresponding decision function is called a classifier. For multi-class classification problem, we use the one-against-one strategy, where a set of binary classifiers are constructed using corresponding data from two classes. While testing, the voting strategy of "Max-Wins" is used to produce the output.

Besides the SVM, we also try the classifier of C4.5 [20], another popular machine learning algorithm. C4.5 is a widely used implementation of the decision tree learning algorithm, in which the learned function is represented by a tree of arbitrary degree that classifies instances [20].

III. EXPERIMENTAL RESULTS

A. *Classification of 4-level beauty*

We randomly select 75% of our 510 facial images as training data set (resulting in 383 training samples), and the rest 25% samples as testing data set(resulting in 127 testing samples). It should be noticed that the number of face samples in different beauty class is not equal. Table Ⅱ gives the detailed information about the size of different beauty classes.

TABLE Ⅱ. THE NUMBER OF SAMPLES IN DIFFERENT CLASSES

	score 0	score 1	score 2	score 3	Total
Training data	45	195	56	87	383
Testing data	11	69	16	31	127

class 0: poor looking, class 1: Common looking, class 2: Good looking, class 3: beautiful looking

The learning of beauty concept is conducted by two machine learning algorithms, SVM and C4.5, with 383 training samples used, and the classification accuracy is performed on 127 testing samples. The experimental results are given in Table Ⅲ.

TABLE Ⅲ. ACCURACY OF CALSSIFICAION OF 4-LEVEL BEAUTY

Classifier / Class #	**SVM** *(Correctly classified number/Total number of this class)*	**C4.5** *(Correctly classified number / Total number of this class)*
3(beautiful)	96.8% (30/31)	90.3% (28/31)
2 (good)	37.5% (6/16)	25.0% (4/16)
1 (common)	84.1% (58/69)	72.5%(50/69)
0 (poor)	45.4% (5/11)	36.4%(4/11)
Average Accuracy	77.9% (99/127)	67.7% (86/127)

From Table Ⅲ, it can be seen that the classification accuracy of 77.9% is obtained for four levels' beauty classification, indicating that the concept of beauty can be automatically learned and classified. Also, the accuracy of SVM is 10.2% higher than C4.5, showing that SVM is more suitable for the problem of facial beauty classification using geometric features. Furthermore, from this table, we found an interesting phenomena that the accuracy of class 3 (beautiful looking) and class 1 (common looking) is significantly higher than the other two classes (good looking and poor looking). This indicates class 0 and class 2 are relatively complex and hard to be learned, they cannot be modeled by simple geometric features.

B. *Classification of beautiful or not*

Inspired from the experiments described above, for the purpose of excluding the uncertain factors in the assessment of class 2 as well as class 0 and further clarifying the efficiency of machine's learning the concept of "beautiful looking", we focus on two classes analysis and design another grading scheme in which we simply score the 510 samples into two classes, beautiful looking (total of 118 faces) and common looking (total of 392 faces). The experimental results are shown in Table Ⅳ.

TABLE Ⅳ. CLASSIFICATION OF BEAUTIFUL OR NOT

Classifier / Class #	**SVM** *(Correctly classified number / Total number)*	**C4.5** *(Correctly classified number / Total number)*
1	96.8% (30/31)	93.5%(293/31)
0	94.8% (91/96)	93.8% (90/96)
Average Accuracy	95.3% (121/127)	93.7%(119/127)

From Table Ⅳ, we can see that for the concept of "beautiful looking" or not, very high classification accuracy (95.3%) is achieved by using SVM, while accuracy of 93.7% can also be obtained by using C4.5. Especially, the classification of "beautiful looking (class 1) " also results in as high accuracy as of 96.8%, which indicates that the concept of ideal beauty can be learned and classified robustly through machine learning approach.

IV. CONCLUSION AND DISCUSSION

This paper presents an automated Chinese female facial beauty classification approach, trained with female facial images and their respective human ratings. The facial images were abstractly represented with 17-element facial geometric feature vectors, and the classification was conducted using SVM and C4.5, the best accuracy we achieved is high up to 95.3%.

In our approach, the data sets adopted has a significantly large size of 510 facial images, and, the beauty extents in these images are widely diversified, the sufficiency and wide diversity of facial image data sets provide ground truth for the training of classifier thus make the experimental results more meaningful and reliable. From the experimental results, we find that though the geometric features we extracted are simple, it proves to be very effective and leads to high accuracy for the recognition of most beautiful level (aesthetical ideal beauty), this finding is surprising and encouraging. On one hand, it is in accordance with the theories regarding the relation between beauty and complexity such as "Low-Complexity Art", "Simplicity lies at the heart of all scientific theories (Occam's razor principle)"[18] . On the other hand, it gives new evidence for the existence of objective quantitative criteria underlying aesthetical ideal beauty. However, the relatively low accuracy of beauty level 2, 1, 0 also indicates that the simple geometric facial features are incomplete for the judgment of other levels of beauty which are relatively more complex, in fact, common sense tells us that there exists other factors such as temperament, familiarity and affection which may exert unignorable influence on the perception of facial beauty in real world.

Overall, the results show that machine can learn the concept of aesthetic ideal facial beauty using supervised machine learning methods. Furthermore, there exists much room for machine to learn facial beauty by means of more complex features other than facial geometrics to cover the real complete perception of facial beauty, which merits our future study.

ACKNOWLEDGMENT

This work is supported in part by NSFC (Grant no. U0735004) and GDSFC (Grant no. 07118074).

REFERENCES

[1]. Aharon I, Etcoff NL, Ariely D, Chabris CF, O'Connor E, Breiter HC, "Beautiful faces have variable reward value: fMRI and behavioral evidence," Neuron 32:537–51, 2001.

[2]. O'Doherty J, Winston J, Critchley H, Perrett D, Burt DM, Dolan RJ, "Beauty in a smile: the role of medial orbitofrontal cortex in facial attractiveness," Neuropsychologia 41:147–55, 2003.

[3]. Etcoff N., "Survival of the Prettiest: The Science of Beauty". New York: Anchor/Doubleday. 1999.

[4]. Alam M, Arndt KA, Dover JS., "Aesthetic surgery: coming of age," Lancet.356:S60. 2000.

[5]. Randy Thornhill and Steven W. Gangestad, "Facial Attractiveness," Trends in Cognitive Sciences, Vol.3, No. 12, December 1999.

[6]. Gillian Rhodes, "The Evolutionary Psychology of Facial Beauty," Annu. Rev. Psychol. 57:199-266, 2006.

[7]. Berry DS., "Attractiveness, attraction, and sexual selection: evolutionary perspectives on the form and function of physical attractiveness." In Advances in Experimental Social Psychology, ed. MP Zanna, 32: 273–342. SanDiego, CA: Academic, 2000.

[8]. PB, Wu C-H., "The Darwin C. 1998/1874. The Descent of Man," Amherst, NY: Prometheu, 1995.

[9]. Cunningham MR, Roberts AR, Barbee AP, Druen PB, Wu C-H. "Their ideas of beauty are, on the whole, the same as ours: consistency and variability in the cross cultural perception of female physical attractiveness," J. Personal. Soc. Psychol. 68:261–79, 1995.

[10]. Larglois JH, Kalakanis L, Rubenstein AJ, Larson A, Hallam M, Smoot M. "Maxims or myths of beauty? A meta-analytic and theoretical review," Psychol. Bull. 126:390–423, 2000.

[11]. Rhodes G, Yoshikawa S, Clark A, Lee K, McKay R, Akamatsu S., "Attractiveness of facial averageness and symmetry in non-Western cultures: in search of biologically based standards of beauty," Perception 30:611–25, 2001.

[12]. Paham Aarabi, D.Hughes, K.Mohajer, M.Emaini, "The automatic measurement of facial beauty," IEEE International Conference on System, Man and Cybernetics, Vol. 4, page(s): 2644-2647, 2001.

[13]. H. Irem, Turkmen Z. Kurt M. Elif KAarsligil, "Global Feature Based Female Facial Beauty Decision System," proceeding of EURASIP07, 2007.

[14]. Y. Eisenthal, G. Dror, E. Ruppin, "Facial Attractiveness: Beauty and the Machine," Neural Computation 18, 119–142, 2006.

[15]. Wen Gao, Bo Cao, Shiguang Shan, Xilin Chen, Delong Zhou, Xiaohua Zhang, Debin Zhao, "The CAS-PEAL Large-Scale Chinese Face Database and Baseline Evaluations," *IEEE Trans. on System Man, and Cybernetics (Part A)*. vol.38, no.1, pp149-161. 2008.1.

[16]. Xiaomei Zhang, Chinese Beauty (in Chinese) Beijing: Xinhua Press 2005.

[17]. Murad Alam, Jeffrey S. Dover, "On beauty Evolution, Psychosocial Considerations, and Surgical Enhancement," ARCH DERMATOL/VOL 137, JUNE, 2001.

[18]. Schmidhuber, J., "Low-Complexity Art," Leonado, Journal of the International Society for the Arts, Sciences, and Technology, 30(2):97-103, MIT Press, 1997.

[19]. V. Vapnik., "Statistical Learning Theory," Springer Verlag, New York, 1998.

[20]. J. R. Quinlan, "C4.5: Programs fro Machine Learning," Morgan Kaufmann, San Francisco, CA, 1993.