Pitch Angle: A Newfound Influential Trait for Image-Based Facial Beauty Perception

Hao Guo, Wenming Han, Fangmei Chen*, and Fuming Sun

Dalian Minzu University, Dalian, Liaoning, 116600, China
*cfm@dlnu.edu.cn  *Corresponding author’s e-mail: fmchen08@126.com

Abstract. One of the most crucial issues in facial beauty study is to find the traits, which determine the perceived attractiveness of a face. During the decades, many traits have been studied extensively, such as proportions, averageness, and symmetry. However, facial beauty seems to be more complex than these traits. In this paper, a face image is represented by a series of interpretable parameters, which are learnt by a CNN based 3DMM fitting algorithm. Based on the SCUT-FBP5500 dataset, we explore the relationship between those parameters and the human rated facial beauty scores. We observe that the pitch angle parameter has a strong correlation with the beauty score. In order to study the causal relationship between the pitch angle and facial beauty, face images with different pitch angles were generated as the stimuli and 39 volunteers were invited to a rating experiment. The results show that the change of the pitch angle has a significant effect on the perceived facial beauty. This finding may explain that head pose plays an important role in the self-portrait shot.

1. Introduction

Facial beauty is an interdisciplinary topic, which has attracted the efforts of researchers from diverse fields such as philosophy, cognitive science, aesthetic plastic surgery, and evolutionary biology[1][2][3][4]. Attractive individuals tend to have better social outcomes, such as to be able to find higher quality jobs [5][6], and to get more opportunities in blind dates [7]. As a result, beauty is an everlasting pursuit of people, and the related industries such as aesthetic plastic surgery, cosmetics, and entertainment are very prosperous.

One of the most crucial questions in facial beauty study is which traits have significant influence on the perceived attractiveness of a face. This is nontrivial. Because not only facial beauty perception is a complex process with multiple semantic concepts, but inter-personal differences exist in subjective judgments of facial beauty [8]. The preference of facial beauty is affected by cultural factors. For example, in China’s Tang Dynasty people prefer the full figure, while in modern times people prefer the slim beauty. Many people in Europe and America like healthy wheat skin color, while most of people in Asian countries tend to pursue light skin. Hence, it is hard to define universal and objective rules of facial beauty perception.

Then can we do some research to deepen the understanding of facial beauty perception, despite the existence of inter-personal differences? The answer is yes. Although personal preferences are not exactly the same, it is found that there is a high degree of agreement on facial beauty perception across ethnicity, social class, age and gender [9][10]. Different kinds of traits, which are assumed to be associated with facial beauty, have been extensively studied during the decades. Averageness[11], symmetry [12], and sexual dimorphism [13] are the most investigated traits in the psychological domain. Classical canons, golden ratios, and anthropometrics are more concerned by the aesthetic
plastic surgery and orthodontics society [14][15]. As the development of image processing techniques, facial landmarks can be automatically located, and the facial proportions can be obtained easily, which promotes the studies based on large scale datasets [16][26]. SCUT-FBP5500 [17]and Beauty799 [18] are such datasets, providing 2D frontal face images and the corresponding human rated beauty scores. Recently, an impressive amount of work is committed to the automatic facial beauty prediction [27]. This is tackled as a supervised [19][20][21][22] or semi-supervised [23][24] learning problem. With the advent of deep learning algorithms, the raw facial image is given as an input to the algorithm, which automatically extracts the putative relevant features in the inference process [25]. The state-of-the-art facial beauty predictors can achieve a correlation about 0.9 with human raters. However, because the training of deep neural networks is a black box process, it is still difficult to interpret what traits make a face beautiful.

In this paper, we learn a parameterized representation for a face image. The parameters have interpretable meanings (e.g., shape, texture, expression, illumination, and pose) and can be used to reconstruct the corresponding 3D face model. Reconstructing 3D face and from 2D images is a hot topic in computer vision [28][29], which is often considered as a model fitting problem. In the past two years, 3D morphable face model (3DMM) based 3D face reconstruction has achieved significant progress, [30]. The by-product is a face image gets a low-dimensional representation in 3DMM subspace, together with the estimated expression, illumination, and pose parameters. For the face images in SCUT-FBP5500 dataset, we learn the parameterized representations and explore the relationship between those parameters and the perceived facial beauty. An interesting finding is that the pitch angle parameter appears to have a strong correlation with the beauty score. In order to verify the effectiveness of this trait, we generated face images with different pitch angles and invited 39 volunteers to judge the attractiveness of them. Statistical results show that the change of the pitch angle has a significant effect on the perceived facial beauty, which supports our finding.

The main contributions of this paper are summarized below:

• We learn a disentangled face representation by a CNN based 3DMM fitting algorithm. In this way, 2D face images are mapped into a low-dimensional space. The meaning of each dimension can be interpreted, and we further examine the association between each dimension and facial beauty. It is the first time to take those parameters into facial beauty study.

• A new finding has been revealed that the pitch angle, which is related to head pose, has a significant effect on the perceived facial beauty based on 2D image. We designed and performed a subjective experiment and verified this finding.

2. Related work

2.1. 3D morphable face model

A 3DMM is a generative model for face shape and appearance [30], which has a long term impact on face analysis studies since it was first proposed by Blanz and Vetter in 1999 [31]. It was trained on 3D face scans that have established dense point-to-point correspondence. Statistical models were learnt by principle component analysis (PCA) on the shape and texture of the 3D faces respectively. 3DMM often serves as a strong prior for monocular 3D face reconstruction. By 3DMM fitting, a 2D face image was mapped into a parametric face space which enabled controlled manipulation, while the external factors such as illumination and camera parameters were disentangled. Currently, the most popular 3DMM was built by merging the Basal Face Model (BFM) [32], which was built with only 200 subjects in neutral expressions, and FaceWarehouse [33] with 150 subjects in 20 different expressions.

2.2. Face representation

Representations of faces have been in focus of vision research for a long time. An important work was the Eigenfaces approach proposed by Sirovich and Kirby [34] and Turk and Pentland [35]. Eigenfaces took the face images in vector form and learnt the statistical model using PCA, with the eigenvectors
representing the main modes of variation. The drawback of Eigenface was that it was limited to a fixed pose and illumination and had no effective representation of shape differences [30]. Another influential work is the Active Appearance Models [36], which combined an explicit shape model and a shape invariant appearance model. These models worked well on a fixed pose and illumination setting. The invention of 3DMMs inspired the idea of face representation in 3DMMs. This representation can disentangle the external factors such as illumination and pose. Analysis-by-synthesis is the commonly used idea to map between the 3D and 2D domain, using 3DMMs as a prior. Compared to AAMs, the fitting of 3DMMs are much more complex. As the development of deep learning, impressive progress has been made on 3DMM fitting [28][37][38][39].

3. Parametric face representation by 3DMM fitting
We adopted the method proposed by [28] to obtain the parametric face representation, which regressed coefficients of a 3DMM face model as well as the illumination and face pose. Specifically, with a 3DMM, the face shape $S$ and the texture $T$ can be represented as follows:

$$S = S(\alpha, \beta) = \bar{S} + B_{id}\alpha + B_{exp}\beta$$

$$T = T(\delta) = \bar{T} + B_{t}\delta$$

where $\bar{S}$ and $\bar{T}$ are the average face shape and average face texture respectively; $B_{id}, B_{exp}, B_{t}$ are the PCA bases of shape, expression and texture respectively; $\alpha, \beta, \delta$ are the corresponding coefficient vectors. Here the BFM model [32] was adopted for the shape and texture, and the expression bases $B_{exp}$ were built from FaceWarehouse [33]. A subset of the bases was selected, resulting in $\alpha \in \mathbb{R}^{80}, \beta \in \mathbb{R}^{64}$, and $\delta \in \mathbb{R}^{80}$.

The fitting algorithm followed the idea of analysis-by-synthesis. Spherical harmonics (i.e. SH) model was used to approximate the lighting information, i.e.,

$$C(n, t | \gamma) = t_i \cdot \sum_{b=0}^{B} \gamma_b \Phi_b(n_i)$$

where $n_i$ is the surface normal. $t_i$ is the skin texture. $\Phi_b: \mathbb{R}^3 \rightarrow \mathbb{R}$ are SH basis functions, and $\gamma_b$ are the corresponding SH coefficients. $B$ was chosen to be 3, such that $\gamma \in \mathbb{R}^{27}$ for 3 channels. The perspective camera model with empirically selected focal length was adopted to realize a 3D to 2D geometric projection, and the 3D face pose is represented by rotation $R \in SO(3)$ and $t \in \mathbb{R}^3$. Finally, there are totally 257 parameters to be estimated, represented by a vector:

$$x = (\alpha, \beta, \delta, \gamma, R, t) \in \mathbb{R}^{257}$$

The order of the parameters is shown in Table 1.

| Dimension range | Attribute          |
|-----------------|--------------------|
| 1-80            | Shape              |
| 81-144          | Expression         |
| 145-224         | Texture            |
| 225-227         | Ruler angle for rotation |
| 228-254         | Illumination       |
| 255-257         | Translation        |

The authors of [28] shared the trained neural network and code of the 3DMM fitting algorithm. Hence, for each input 2D face image, we can obtain its 257-dimensional representation and the reconstructed 3D face model. The flowchart of generating 3D face models using 2D face images is
shown in Figure 1. Furthermore, by changing the coefficients, we can manipulate the generated 3D face.

4. Association study between face parameters and beauty scores
Using the method above, we can obtain a compact parametric representation for each face image. It is interesting to investigate which parameters are related to facial beauty.

4.1. Dataset
SCUT-FBP5500 is a public dataset [17] for facial beauty analysis, which contains 5500 frontal faces aged from 15 to 60 with neutral expression. It is divided into four subsets: 2000 Asian females, 2000 Asian males, 750 Caucasian females and 750 Caucasian males. All the images were labelled by 60 persons with a 5-point scale. The mean opinion score (MOS) is used to measure the facial beauty.

4.2. Correlation analysis
Based on the SCUT-FBP5500 dataset, we calculate Pearson’s correlation coefficient:

\[
 r(X,Y) = \frac{\text{Cov}(X,Y)}{\sqrt{\text{Var}[X] \text{Var}[Y]}}
\]  

(4)

where \( \text{Cov}(X,Y) \) is the covariance of \( X \) and \( Y \), \( \text{Var}[X] \) is the variance of \( X \), and \( \text{Var}[Y] \) is the variance of \( Y \). In our case, each dimension of the 257-dimensional parameters of the 5500 samples is regarded as \( X \), i.e., a factor to be examined, and the vector of the corresponding MOSs is regarded as \( Y \). The absolute value of \( r \) can be used to measure the association strength. Figure 2 plots \(|r|\) of each facial parameter, and the top 5 parameters with the largest \(|r|\) is shown in Table 2. We can see that parameters with large correlations are mostly in the texture and shape categories, which is reasonable because those parameters determine the identity. However, an exception is that the second one is belongs to the pose category, specifically, the pitch angle parameter.
5. Verify the effectiveness of the pitch angle trait
The strong correlation between the pitch angle and facial beauty arouses our curiosity. Is the effect of the pitch angle just a coincidence on this dataset, or is it really an influential factor for facial beauty perception? To answer this question, we need an intervention experiment, i.e., generating face images with different pitch angles and inviting human raters to compare the attractiveness of them. If the outcomes differ, we say that the pitch angle trait has a causal effect on facial beauty.

5.1. Stimuli generation
From the SCUT-FBP5500 dataset, we randomly selected 3 faces from each of the 4 categories, i.e., Asian females, Asian males, Caucasian females, and Caucasian males. From each of the faces, $I_{\text{pitch}0}$, we changed its pitch angle using the MeshLab software and generated two new faces, $I_{\text{pitch}+}$ and $I_{\text{pitch}-}$, where “pitch+” means increasing the pitch angle slightly and “pitch-” means decreasing the pitch angle slightly. The sketch of the pitch angle on a head model is shown in Figure 3. The generated face images are shown in Figure 4, which were used as stimuli in the subjective experiment. Hence, there are totally 12 groups of stimuli. Each group includes 3 images, $I_{\text{pitch}+}$, $I_{\text{pitch}0}$, and $I_{\text{pitch}-}$.
5.2. Procedures
The face images were displayed group by group and side by side in a questionnaire. For each group, the positions of the 3 images are randomized. Totally 39 participants (aged between 20 and 25) were invited to make forced-choice comparisons among the 3 face images in each group, i.e. they had to choose the one they thought more attractive than the other two. None of these participants was familiar to the faces in the training data. They then submitted their choices directly through the web.

5.3. Experimental results
Figure 5 shows the number of participants who chose $I_{\text{pitch-}}$, $I_{\text{pitch0}}$, or $I_{\text{pitch+}}$ as the more attractive one. We can see that the percentage of choices has the relationship $I_{\text{pitch-}} > I_{\text{pitch0}} > I_{\text{pitch+}}$ in general for most stimulus faces. In other words, for frontal face images, slightly decreasing the pitch angle is more likely to increase the attractiveness.

We then considered three different hypotheses, i.e. (a) $H_0: f(I_{\text{pitch-}}) \leq f(I_{\text{pitch0}})$; $H_1: f(I_{\text{pitch-}}) > f(I_{\text{pitch0}})$, (b) $H_0: f(I_{\text{pitch0}}) \leq f(I_{\text{pitch+}})$; $H_1: f(I_{\text{pitch0}}) > f(I_{\text{pitch+}})$, and (c) $H_0: f(I_{\text{pitch-}}) \leq f(I_{\text{pitch+}})$; $H_1: f(I_{\text{pitch-}}) > f(I_{\text{pitch+}})$. The null hypotheses in all these three cases claim that decreasing the pitch angle of a frontal face cannot improve its attractiveness, and the alternative hypotheses support the hypothesis: decreasing the pitch angle of a frontal face will improve its attractiveness. The p-value over the 12 groups of stimulus faces under these cases are shown in Table 3.

For each case, paired sample $t$ test was performed. The statistic $t$ is calculated as follows:

$$t = \frac{X_D - \mu_0}{s_D / \sqrt{n}},$$

where $X_D$ is the difference between the means of the two groups, $\mu_0$ is the hypothesized mean difference, $s_D$ is the standard deviation of the differences, and $n$ is the sample size.
where $\overline{D}$ and $s_D$ are the average and standard deviation of the differences between all pairs. The constant $\mu_0$ is zero. $n$ represents the number of pairs, which is 12 in our experiment. The degree of freedom used is $n - 1$. The results are shown in Table 3. We can see that at a significant level of 0.05, the $H_0$ hypotheses were rejected in all the three cases, which means that in statistical sense, decreasing the pitch angle of a face will improve its attractiveness (based on 2D face images).

Figure 5. Result of the group-wise voting experiment

Table 3. Results of the three hypothesis cases (significant level 0.05).

| Case | Statistic t | p-value | Decision |
|------|-------------|---------|----------|
| (a)  | 7.891       | 7.441e-6| Reject $H_0$ |
| (b)  | 4.946       | 4.386e-4| Reject $H_0$ |
| (c)  | 15.717      | 6.962e-9| Reject $H_0$ |

6. Conclusions

This paper performs disentangled face representation based on a 3DMM fitting algorithm. Then based on SCUT-FBP5500 dataset, we study the association between different face parameters and facial beauty scores. We find that the pitch angle trait has a strong correlation with the facial beauty and design subjective experiment to verify this finding. This finding reveals that pose is an important factor for 2D face aesthetic. This study is based on the SCUT-FBP5500 dataset, which includes only frontal faces. The variations of yaw and roll angles are very limited. Although no obvious correlation is found between these angles and facial beauty in this study, it is worthwhile to pay attention to them when studying facial beauty with a more diversified dataset.

Acknowledgments

We would like to thank the editor and the anonymous reviewers for their insightful comments and suggestions. We are also grateful to all the volunteers who participated in the experiments. This work was supported by National Natural Science Foundation of China under Grant 61702082, Grand 61702081, Grand 61972068, and Grand 61976042.

References

[1] Vučinić, N., Tubbs, R.S., Erić, M., Vujić, Z., Marić, D. (2020) What do we find attractive about the face?: Survey study with application to aesthetic surgery. Clin. Anat., 33: 214–222.
[2] Eggerstedt, M., Rhee J., Urban M. et al. (2020) Beauty is in the eye of the follower: Facial aesthetics in the age of social media. Am. J. Otolarngol. Head Neck Med. Surg., 41: 102643.
[3] Han, S., Liu, S., Gan Y. et al. (2020) Repeated exposure makes attractive faces more attractive: Neural responses in facial attractiveness judgement. Neuropsychologia, 139: 107365.
[4] Little, A.C., Jones, B.C., Debruine, L.M. (2011) Facial attractiveness: evolutionary based research. Philosophical Transactions of the Royal Society of London, 366(1571): 1638-1659.
[5] Sala E., Terraneo M., Lucchini M., Knies G. (2013) Exploring the impact of male and female facial attractiveness on occupational prestige. Research in Social Stratification and Mobility, 31: 69-81.

[6] Peng, L., Cui, G., Chung, Y., Zheng, W. (2020) The faces of success: Beauty and ugliness premiums in e-commerce platforms. J. Mark, 84: 67–85.

[7] Walster E., Aronson V., Abrahams D., Rottman L. (1966) Importance of physical attractiveness in dating behavior. Journal of Personality and Social Psychology, 4(5): 508-516.

[8] Ibáñez-Berganza, M., Amico, A., Loreto, V. (2019) Subjectivity and complexity of facial attractiveness. Sci. Rep, 9: 1–12.

[9] Cunningham M.R., Roberts A.R., Barbee A.P., Druen P.B., Wu C.H. (1995) “Their ideas of beauty are, on the whole, the same as ours”: Consistency and variability in the cross-cultural perception of female physical attractiveness. Journal of Personality and Social Psychology, 68(2): 261-279.

[10] Langlois J.H., Kalakanis L., Rubenstein A.J., Larson A., Hallam M., Smoot M. (2000) Maxims or myths of beauty? A meta-analytic and theoretical review. Psychological Bulletin, 126(3): 390–423.

[11] Perrett D., May K., Yoshikawa S. (1994) Facial shape and judgements of female attractiveness. Nature, 368(6468): 239–242.

[12] Perrett D., Burt D., Penton-Voak I., Lee K., Rowland D., Edwards R. (1999) Symmetry and human facial attractiveness. Evolution Human Behavior, 20(5): 295–307.

[13] Perrett D. et al. (1998) Effects of sexual dimorphism on facial attractiveness. Nature, 394(6696): 884–887.

[14] Ulrich, L., Dugelay, J.L., Vezzetti, E., Moos, S., Marcolin, F. (2020) Perspective morphometric criteria for facial beauty and proportion assessment. Appl. Sci., 10.

[15] Samizadeh, S. (2019) The ideals of facial beauty among Chinese aesthetic practitioners: Results from a large national survey. Aesthetic Plast. Surg., 43: 102–114.

[16] Chen, F., Zhang, D. (2014) Evaluation of the Putative Ratio Rules for Facial Beauty Indexing. In: 2014 Int. Conf. Med. Biometrics, Shenzhen, pp: 181–188.

[17] Liang, L., Lin, L., Jin, L., Xie, D., Li, M. (2018) SCUT-FBP5500: A Diverse Benchmark Dataset for Multi-Paradigm Facial Beauty Prediction. In: Int. Conf. Pattern Recognit. 2018, Augus, pp. 1598–1603.

[18] Zhang, D., Chen, F., Xu, Y. (2016) Computer Models for Facial Beauty Analysis. Springer International Publishing, Switzerland.

[19] Eisenthal Y., Dror G., Ruppin E. (2006) Facial attractiveness: Beauty and the machine. Neural Comput., 18(1): 119–142.

[20] Gunes H., Piccardi M. (2006) Assessing facial beauty through proportion analysis by image processing and supervised learning. Int. J. Human-Comput. Studies, 64(12): 1184–1199.

[21] Schmid K., Marx D., Samal A. (2008) Computation of a face attractiveness index based on neoclassical canons, symmetry, and golden ratios. Pattern Recognition, 41(8): 2710–2717.

[22] Fan J., Chau K.P., Wan X., Zhai L., Lau E. (2012) Prediction of facial attractiveness from facial proportions. Pattern Recognition, 45(6): 2326–2334.

[23] Zhang L., Zhang D., Sun M., Chen, F. (2017) Facial beauty analysis based on geometric feature: Toward attractiveness assessment application. Expert Syst. Appl., 82: 252–265.

[24] Dornaika F., Wang K., Arganda-Carreras I., Elorza A., Moujahid A. (2020) Toward graph-based semi-supervised face beauty prediction. Expert Syst. Appl., 142: 112990.

[25] Zhai Y. et al. (2020) Asian female facial beauty prediction using deep neural networks via transfer learning and multi-channel feature fusion. IEEE Access, 8: 56892–56907.

[26] Zhang D., Zhao Q., Chen, F. (2011) Quantitative analysis of human facial beauty using geometric features. Pattern Recognition, 44: 940–950.

[27] Chen, F., Xiao, X., Zhang, D. (2018) Data-driven facial beauty analysis: Prediction, retrieval and manipulation. IEEE Trans. Affect. Comput., 9: 205–216.
[28] Deng, Y. et al. (2019) Accurate 3D Face Reconstruction with Weakly-Supervised Learning: From Single Image to Image Set. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops.

[29] Gao Z., Zhang J., Guo Y., Ma C., Zhai G., Yang X. (2020) Semi-Supervised 3D Face Representation Learning From Unconstrained Photo Collections. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, pp. 348-349.

[30] Egger B., Smith W.A., Tewari A. et al. (2020) 3D Morphable Face models - past, present, and future. ACM Transactions on Graphics (TOG), 39(5): 1-38.

[31] Blanz V., Vetter T. (1999) A Morphable Model for the Synthesis of 3D Faces. In: Proc. 26th Annu. Conf. Comput. Graph. Interact. Tech. SIGGRAPH ’99. pp. 187-194.

[32] Paysan P., Knothe R., Amberg B., Romdhani S., Vetter, T. (2009) A 3D Face Model for Pose and Illumination Invariant Face Recognition. In: 6th IEEE Int. Conf. Adv. Video Signal Based Surveillance, AVSS 2009, pp. 296–301.

[33] Cao C., Weng Y., Zhou S., Tong Y., Zhou, K. (2014) FaceWarehouse: A 3D facial expression database for visual computing. IEEE Trans. Vis. Comput. Graph, 20: 413–425.

[34] Sirovich L., Kirby M. (1987) Low-dimensional procedure for the characterization of human faces. Journal of the Optical Society of America A, 4(3): 519–524.

[35] Turk M., Pentland A. (1991) Eigenfaces for recognition. Journal of Cognitive Neuroscience, 3(1): 71–86.

[36] Cootes T.F., Edwards G.J., Christopher J.T. (1998) Active Appearance Models. In: Proc. European Conference on Computer Vision.

[37] Guo Y., Zhang J.Z., Cai J., Jiang B., Zheng J. (2019) Cnn-based real-time dense face reconstruction with inverse- rendered photo-realistic face images. IEEE Transactions on Pattern Analysis and Machine Intelligence, 41(6): 1294-1307.

[38] Tewari A., Zollhöfer M., Kim H. et al. (2017) MoFa: Model-Based Deep Convolutional Face Autoencoder for Unsupervised Monocular Reconstruction. In: International Conference on Computer Vision (ICCV), pp. 1274-1283.

[39] Tran A.T., Hassner T., Masi I., Medioni G. (2017) Regressing Robust and Discriminative 3d Morphable Models with a very Deep Neural Network. In IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1493–1502.