Exploring Relationship Between Face and Trustworthy Impression Using Mid-level Facial Features

Yan Yan[1], Jie Nie[2], Lei Huang[1], Zhen Li[1], Qinglei Cao[1], and Zhiqiang Wei[1](✉)

[1] College of Information Science and Engineering, Ocean University of China, Qingdao China
{yanyan.azj,ithuanglei,lizhen0130,cql.levi}@gmail.com,
weizhiqiang@ouc.edu.cn

[2] Department of Computer Science and Technology, Tsinghua University, Beijing China
niejie@tsinghua.edu.cn

Abstract. When people look at a face, they always build an affective subconscious impression of the person which is very useful information in social contact. Exploring relationship between facial appearance in portraits and personality impression is an interesting and challenging issue in multimedia area. In this paper, a novel method which can build relationship between facial appearance and personality impression is proposed. Low-level visual features are extracted on the defined face regions designed from psychology at first. Then, to alleviate the semantic gap between the low-level features and high-level affective features, mid-level feature set are built through clustering method. Finally, classification model is trained using our dataset. Comprehensive experiments demonstrate the effectiveness of our method by improving 26.24 % in F1-measure and 54.28 % in recall under similar precision comparing to state-of-the-art works. Evaluation of different mid-level feature combinations further illustrates the promising of the proposed method.

Keywords: Portrait · Affective subconscious impression · Personality · Mid-level feature

1 Introduction

People could build impression including facial appearance and affective subconscious impression at first sight. The facial appearance is respect to the person's facial attributes, such as her/his eye size, face shape, etc. The affective subconscious includes personality impression primarily, e.g., "he is reliable". Many classical and famous psychology works have find that facial appearance correlated with impression of personality. How to infer relationship between face appearance and personality impression becomes an interesting and challenging problem in multimedia area, which can predict others' personality by 'reading' their faces in artificial intelligence.

Some works in recent have focused on exploring relationship between visual patterns and personality traits. Richard et al. have done some works to uncover the information driving first impression of social traits judgments, such as trustworthiness or dominance, using an attribute-based approach [1]. They create a quantitative model that can predict

Q. Tian et al. (Eds.): MMM 2016, Part I, LNCS 9516, pp. 540–549, 2016.
DOI: 10.1007/978-3-319-27671-7_45

first impressions of previously unseen ambient images of faces from a linear combination of facial attributes. This study shows that despite enormous variation in ambient images of faces, a substantial proportion of the variance in first impressions can be accounted for through linear changes in objectively defined features. Nie et al. presented a method to infer personality impression from portrait [2]. They used features including background, photograph color, person gender in portraits and focused on four personality impression of the whole portrait. In our work, we focus on face area. Because the appearance of face area is almost the same during ones' whole life, but clothes, photograph background and so on are changeful.

In our previous work, we have found that using facial features can predict people's trustworthiness impression at first sight [3]. From a psychological perspective, we propose novel personality-toward features combining eleven permanent facial features and five transient facial features. All sixteen features are extracted from five main facial features, consisting of eyebrow, eye, nose, mouth and face shape. For different facial regions, we used different feature extraction methods, for example, Histogram of Gradients (HOG) is used to describe eyebrow shape, and Euclidean Distance (ED) is used to describe width of eyes. This work has provided theoretical and empirical insight into finding relationship between facial appearance and trustworthiness impression. In this work, we will further explore the relationship and consider more factors of extracted features in previous work. For example, the dimension of different face regions descriptors are quite different (e.g. Scale-invariant Feature Transform (SIFT) descriptors has 127 dimensions, ED descriptors just has one dimension). We will settle the problem of vastly different dimensions of each facial trait and semantic gap between the low-level features and affective high-level features.

In this paper, we build a framework shown in Fig. 1 through proposing a new mid-level feature set from low-level features using unsupervised learning method to find relationship between facial appearance and personality impression. The mid-level feature set with mid-semantic meaning contains important discriminative information of each face region and maybe effectiveness for predicting personality impression. Relationship between face features and personality impression is inferred by the mid-level facial feature set through supervised learning method. First, we extract the low-level features from face region as face local descriptors. Then we use unsupervised feature learning algorithms translate local descriptors into a set of mid-level feature, which can represent the face in portrait. After that, Support Vector Machine (SVM) [4] is used to find relationship between face features and personality impression. In experiment section, we use K-fold cross-validation to verify the effective of our data set and then compare the effectiveness of our method with other methods. Multinomial logistic regression and Likelihood ratio test are used to compare the contribution of different mid-level feature combination. Performance of clustering low-level features into mid-level feature and correlated face colleges are demonstrated.

The remaining of the paper is organized as follows. Related works are introduced in Sect. 2. The feature extraction method is described in Sect. 3. In Sect. 4, comprehensive experiments are done and comparative results are given. Finally, some conclusions are drawn in Sect. 5.

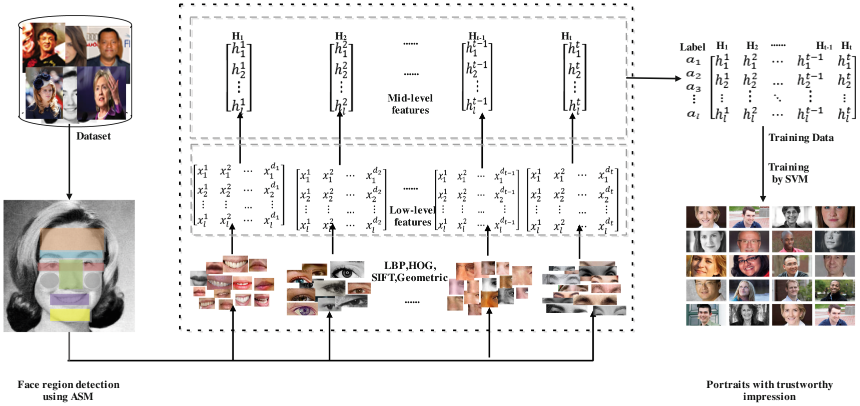

Fig. 1. Framework of the proposed method

2 Related Works

There are numerous works in current focusing on semantic learning in multimedia area using describable visual attributes. The describable visual attributes are regarded as labels that can be given to an image to describe its appearance and widely used to solve classic problems in computer vision area. Kumar et al. have introduced the usage of describable visual attributes, examples of face attributes include gender, age, jaw shape, nose size, etc., for face verification and image search [5–7]. Lin et al. tried to find the Latent Human Topic (LHT) in facial attributes comparing to Kumar et al.'s work [8]. [9, 10] have shown how mid-level semantic attributes can be used synergistically with low-level features for both identification and re-identification. These visual attribute labels in these works were learned through manual annotation, which meaning that all visual attributes are built on personal perception. Cao et al. proposed an adaptive system that allows for ties when collecting ranking data, in which each image is represented by a set of facial attributes [11]. [12, 13] focused on automatic facial attribute detection by different methods. Most of these works above, focused on extracting mid-level which related to attire, biometrics, appendix, but not natural face attributes.

Face is a distinctive region, which has settled region structure, e.g. symmetric regions, two brows located upon two eyes, one nose under middle of eyes. They give people different impression from different angles and change with different facial expressions, which can make a different impression comparing to their real facial attributes. In our work, we use a new mid-level feature set which is clustered from low-level features through unsupervised learning method to describe natural attribute of a face. Many works have applied cluster analysis method into feature classification using unsupervised learning. Sheikh et al. presented a new approach that recognizes facial expressions automatically and also to show the effectual outcome of their approach using K-Means Clustering [14].

3 Proposed Method

In this section, we extract low-level features from different face regions and cluster them into a vector of mid-level features to build a vector of facial features which is personality-toward.

3.1 Low-level Features

In this paper, we focused our work on main regions of face area. We extracted eighteen type features from main regions e.g. eye, brow, nose, mouth and relationship between regions using methods like HOG, Local Binary Patterns (LBP), SIFT and geometric descriptor [15]. Following Fig. 2 shows the workflow of low-level features extraction and description.

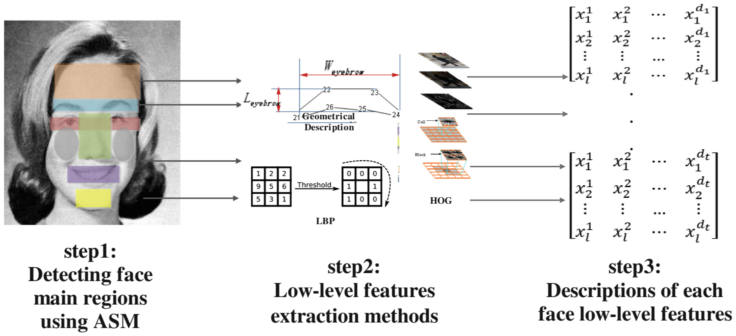

Fig. 2. Workflow of low-level features extraction and description

At first, we utilize Active Shape Model (ASM) [16] to detect main facial main regions and obtained the contours of them. Then the face area was normalized into 128 × 128 pixels, and converted into gray image to avert influence of illumination intensity. At last, a mixture of features descriptors consisting of HOG, LBP, SIFT and geometrical descriptions were adopted to describe personality-toward features. From each person image, we first extract eighteen kinds of face features which add up to 3412-dimensional low-level features.

3.2 Mid-level Features

As mentioned in the Sect. 3.1, each face is represented as eighteen kinds of face features which add up to 3412-dimensional low-level features. Conventionally, this extremely high dimensional vector is directly fed to an SVM for classification. However, the dimension of each attribute is quite different extracted by different methods. As result, performing classification in such feature space is susceptible to overfitting. Furthermore, each attribute may make important contribution to classification. Hence, we use a vector that contains eighteen dimension face features representing a face, in which each dimension clustered from the corresponding low-level features. As one of the most popular and simple clustering algorithms, K-means algorithms [17] is used to cluster the corresponding low-level features, whose result is different from its visual attribute and we call it Mid-level feature.

For each facial portrait, suppose there are T mid-level features, we denote sample data as $X_t = [x_{t,}^1, x_{t,}^2, x_{t,}^3 \ldots, x_{t,}^{d_t}]$, where $t = 1, 2, \ldots, \mathrm{T}$, meaning the t-th mind-level feature, d_t represents the dimension of t-th low-level feature. Then each low-level feature set X_t clustered into a single feature vector H_t valuing $C_t = \{k, k = \{1, 2, \{inK\ \}\}$, K is the number of features status (e.g. the number of eye status is 3 including open, close and slightly open).

For each face, we use all the mid-level features $\mathrm{Z} = [H_1, H_2, \ldots, H_T]$ as training data and associate a training label $y_i \in [0, 1]$, where 0 means trustworthy and 1 means untrustworthy. SVM is used to classify trustworthy personality using mid-level features clustering. We select the RBF as kernel.

4 Experiments

In this section, we conduct comprehensive evaluations of our method. The dataset is described first. Then, experiments are performed to evaluate the proposed approach. We divide our experiments into three parts. First, we use K-fold (K = 5) cross-validation method to draw statistically convincing conclusion and compare our method with state-of-the-art method. Second, we analyze combinations of different number mid-level features contributions to trustworthy impression. Finally, a collage of portraits and clustering results are provided to demonstrate the mid-level features qualitatively.

4.1 Dataset

We use a challenging dataset in work of Ref. [3], which contains 2010 portraits and be labeled the ground-truth by about 250 volunteers. For each portrait, the volunteers were asked to label a tag about trustworthiness or untrustworthiness on it at their first impression. If the consistence of trustworthy by different raters achieved 70 %, the image was tagged with 'trustworthy'.

1404 images were randomly chosen as the training set and the other 606 images were used as the testing set. The details of the training set and testing set are shown in Table 1.

Table 1. Sample number of training set and testing set

Group	Training set		Testing set		Total
	Trustworthy	Untrustworthy	Trustworthy	Untrustworthy	
Portraits number	522	882	273	333	2010

4.2 Experiments and Discussions

Comparison with Other Methods. First we compare the classification accuracy of our new method with Ref. [3]. In order to draw statistically convincing conclusions, it is important to estimate the uncertainty of such estimates. In this experiment, we used K-fold (K = 5) cross-validation method, which can make it possible to reproduce the results and be unbiased to a specific data distribution.

Table 2 lists the classification results of our method and Ref. [3] using three evaluation metrics, i.e., precision, recall and F1- measure. We use the average accuracy of K-fold cross-validation method as the result. From Table 2, we can find that our method has significantly outperformed method in Ref. [3], which gains a great improvement of 54.28 % in recall with a small decrease of precision (6.14 %). Further, our method gains an improvement of 26.24 % in F1-measure.

Table 2. Performance comparison in F1-measure

Methods	Precision	Recall	F1-measure
Reference [3] method	71.10 %	37.00 %	48.70 %
Our method	64.96 %	91.28 %	74.94 %

The improvement benefits from our proposed feature set clustered from low-level features since the feature set is mid-level that can alleviate the semantic gap compared to low-level features. Moreover, feature set can alleviate the influence of the different dimensions of each low-level feature. The semantic meanings of mid-level feature set will be highlighted using qualitative experiment in Fig. 4.

To verify effectiveness of our mid-level features, we use Multinomial logistic regression model to fit the data and likelihood ratio test to demonstrate the contribution of each mid-level feature. Results are shown in following Fig. 3.

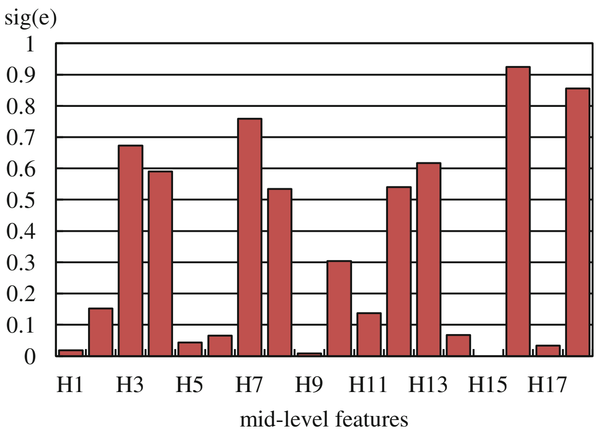

Fig. 3. Likelihood ratio test of each mid-level feature

Inferring from experience, we set a threshold as 0.05. If the likelihood of the mid-level feature is lower than threshold, this mid-level feature is correlation with personality. From Fig. 3 we can see that the likelihood of mid-level features $\{H_1, H_5, H_9, H_{15}, H_{17}\}$ are lower than 0.05, which means that these five mid-level feature are correlation with trustworthy impression.

To further represent the correlation of mid-level features and trustworthy impression, we exhibit the precision, recall and F1-measure of different number mid-level features in Table 3.

Table 3. Precision, recall and F1-measure of different number mid-level feature combination. The row of number means the combination of different number mid-level features determined by the value of likelihood ratio test. For example, combination of three mid-level features is $\{H_9, H_{15}, H_{17}\}$, which are the lowest three likelihood ratio test features. High performance is in bold.

Number	3	5	7	9	11	13	15	18
Precision	**47.37**	45.83	45.07	53.70	44.29	43.48	44.29	43.48
Recall	**72.00**	66.00	64.00	58.00	62.00	60.00	62.00	60.00
F1-measure	**57.14**	54.09	52.89	55.76	51.66	50.42	51.66	50.42

From Table 3, we can see that combination of three mid-level features perform best in all combinations, which means that we can use only three mid-level features to judge impression of a face is trustworthy or not. From the table we also can find that the combination of nine mid-level features performs better than other combinations in F1-measure and precision except for three mid-level features. It indicates that face features are not independent, such as facial expression is determined by several face AUs (Action Units).

Qualitative Demonstration of Mid-level Features. In this subsection, we demonstrate the results of clustering low-level features into mid-level features and corresponding face collages in Fig. 4. In this figure, we demonstrate four representative mid-level features. From the figure, we can see that there is common sense of each mid-level feature. Figure 4(a) is features related with eye, we can see that eyes in the first cluster are bigger than second. The smaller include inborn and squinted eyes leaded by smile or confused expressions. Figure 4(b) is features related with distance of eye and brow, we can see that distance in the first cluster are more far than second. Far distance maybe produced by exaggerated facial expressions. Figure 4(c) is features related with jaw region, we can see that jaws in the first cluster are bigger than second. This face trait could not change with facial expression or other facial movement. But they can form different impression through different angles. Figure 4(d) is features related with mouth station, we can see that mouths in the first cluster are tight close, second cluster are slight open and third cluster are wide open. This face trait has nothing to inborn feature, but just be changed by lips' movement. We also can see that open mouths appear with facial expression of laugh and surprise.

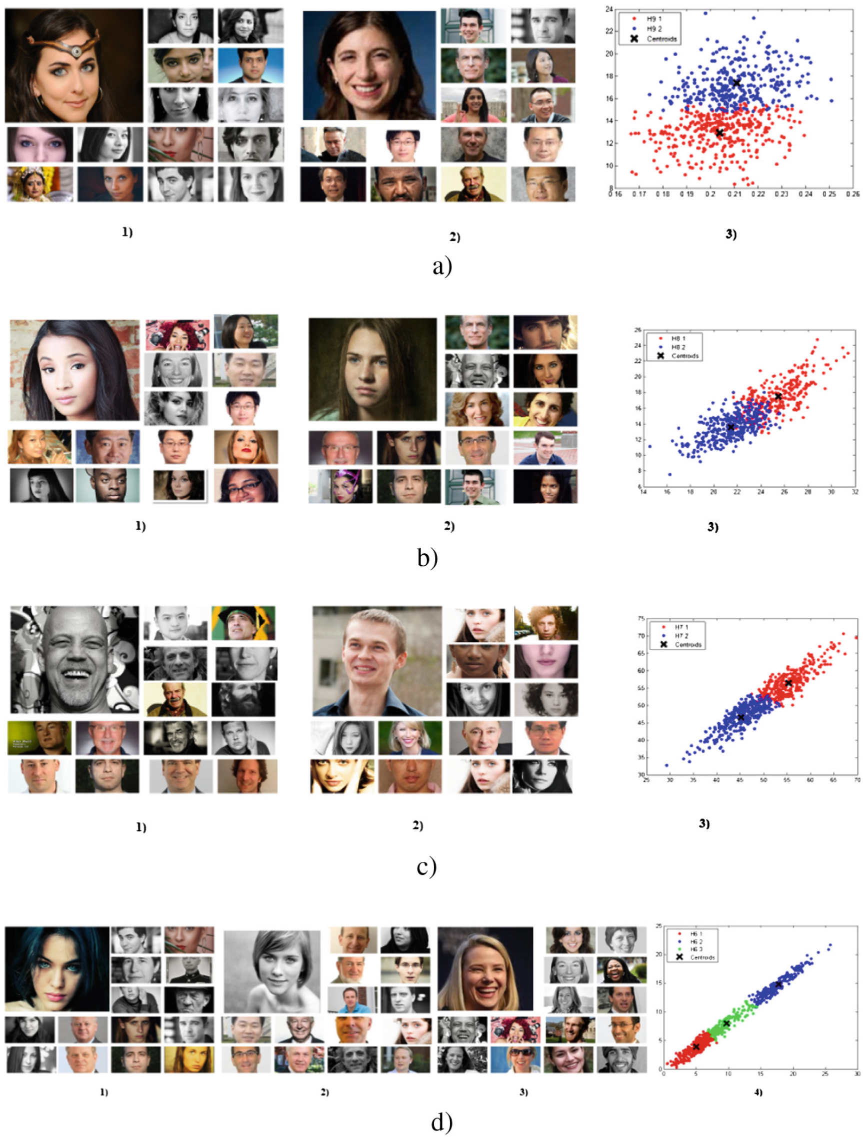

Fig. 4. Demonstrations of four representative mid-level features and corresponding face collages. (a) Represent eye region mid-level feature H_9, (b) Represent region between eye and brow mid-level feature H_8, (c) Represent jaw region mid-level feature H_7, (d) Represent mouth region mid-level feature H_{17}. (a)-3, (b)-3, (c)-3, (d)-4 show result of clustering low-level features into mid-level feature.

5 Conclusions

In this work, we proposed a new mid-level feature set which built from low-level features using clustering method to find the relationship between facial appearance and personality impression. The mid-level feature set with mid-semantic meaning contains important discriminative information of each face region that can alleviate the semantic gap between the low-level features and affective high-level features and make impression analysis more effective. Experimental results demonstrate that our method can get promising performance. Future work will still focus on this interesting and challenging topic of personality analysis from portraits.

Acknowledgments. This work is supported by the National Nature Science Foundation of China (No. 61402428, No. 61202208); the Fundamental Research Funds for the Central Universities (No. 201413021, No. 201513016).

References

1. Yan, Y., Nie, J., Huang, L., Li, Z., Cao, Q., Wei, Z.: Is your first impression reliable? trustworthy analysis using facial traits in portraits. In: He, X., Luo, S., Tao, D., Xu, C., Yang, J., Hasan, M.A. (eds.) MMM 2015, Part II. LNCS, vol. 8936, pp. 148–158. Springer, Heidelberg (2015)
2. Kumar, N., Berg, A.C., Belhumeur, P.N., Nayar, S.K.: Attribute and simile classifiers for face verification. In: 2009 IEEE 12th International Conference on Computer Vision, pp. 365–372. IEEE, September 2009
3. Lin, C.H., Chen, Y.Y., Chen, B.C., Hou, Y.L., Hsu, W.: Facial attribute space compression by latent human topic discovery. In: Proceedings of the ACM International Conference on Multimedia, pp. 1157–1160. ACM, November 2014
4. Layne, R., Hospedales, T.M., Gong, S.: Towards person identification and re-identification with attributes. In: Fusiello, A., Murino, V., Cucchiara, R. (eds.) ECCV 2012 Ws/Demos, Part I. LNCS, vol. 7583, pp. 402–412. Springer, Heidelberg (2012)
5. Klare, B.F., Klum, S., Klontz, J.C., Taborsky, E., Akgul, T., Jain, A.K.: Suspect identification based on descriptive facial attributes. In: 2014 IEEE International Joint Conference on Biometrics (IJCB), pp. 1–8. IEEE, September 2014
6. Cao, C., Kwak, I.S., Belongie, S., Kriegman, D., Ai, H.: Adaptive ranking of facial attractiveness. In: 2014 IEEE International Conference on Multimedia and Expo (ICME), pp. 1–6. IEEE, July 2014
7. Celli, F., Bruni, E., Lepri, B.: Automatic personality and interaction style recognition from facebook profile pictures. In: Proceedings of the ACM International Conference on Multimedia, pp. 1101–1104. ACM, November 2014
8. Datta, A., Feris, R., Vaquero, D.: Hierarchical ranking of facial attributes. In: 2011 IEEE International Conference on Automatic Face & Gesture Recognition and Workshops (FG 2011), pp. 36–42. IEEE, March 2011
9. Sheikh, T., Agrawal, S.: Performance comparison of an effectual approach with k-means clustering algorithm for the recognition of facial expressions. Int. J. Comput. Sci. Mob. Comput. **2**(6), 300–306 (2013)
10. Tian, Y.L., Kanade, T., Cohn, J.F.: Recognizing action units for facial expression analysis. IEEE Trans. Pattern Anal. Mach. Intell. **23**(2), 97–115 (2001)

11. Cootes, T.F., Taylor, C.J., Cooper, D.H., Graham, J.: Active shape models-their training and application. Comput. Vis. Image Underst. **61**(1), 38–59 (1995)
12. Nie, J., Cui, P., Yan, Y., Huang, L., Li, Z., Wei, Z.: How your portrait impresses people?: inferring personality impressions from portrait contents. In: Proceedings of the ACM International Conference on Multimedia, pp. 905–908. ACM, November 2014
13. Vernon, R.J., Sutherland, C.A., Young, A.W., Hartley, T.: Modeling first impressions from highly variable facial images. Proc. Natl. Acad. Sci. **111**(32), E3353–E3361 (2014)
14. Chang, C.C., Lin, C.J.: LIBSVM: a library for support vector machines. ACM Trans. Intell. Syst. Technol. (TIST) **2**(3), 27 (2011)
15. Kumar, N., Berg, A.C., Belhumeur, P.N., Nayar, S.K.: Describable visual attributes for face verification and image search. IEEE Trans. Pattern Anal. Mach. Intell. **33**(10), 1962–1977 (2011)
16. Kumar, N., Belhumeur, P.N., Nayar, S.K.: FaceTracer: a search engine for large collections of images with faces. In: Forsyth, D., Torr, P., Zisserman, A. (eds.) ECCV 2008, Part IV. LNCS, vol. 5305, pp. 340–353. Springer, Heidelberg (2008)
17. Jain, A.K.: Data clustering: 50 years beyond K-means. Pattern Recogn. Lett. **31**(8), 651–666 (2010)