Do Others Perceive You As You Want Them To? Modeling Personality based on Selfies

Sharath Chandra Guntuku
SCE
NTU, Singapore
sharathc001@e.ntu.edu.sg

Lin Qiu
HSS
NTU, Singapore
linqiu@ntu.edu.sg

Sujoy Roy
I^2R
A*STAR, Singapore
sujoy@i2r.a-star.edu.sg

Weisi Lin
SCE
NTU, Singapore
wslin@ntu.edu.sg

Vinit Jakhetiya
ECE
HKUST, Hong Kong
vjakhetiya@ust.hk

ABSTRACT

In this work, selfies (self-portrait images) of users are used to computationally predict and understand their personality. For users to convey a certain impression with selfie, and for the observers to build a certain impression about the users, many visual cues play a significant role. It is interesting to analyse what these cues are and how they influence our understanding of personality profiles. Selfies of users (from a popular microblogging site, Sina Weibo) were annotated with mid-level cues (such as presence of duckface, if the user is alone, emotional positivity etc.) relevant to portraits (especially selfies). Low-level visual features were used to train models to detect these mid-level cues, which are then used to predict users' personality (based on Five Factor Model). The mid-level cue detectors are seen to outperform state-of-the-art features for most traits. Using the trained computational models, we then present several insights on how selfies reflect their owners' personality and how users' are judged by others based on their selfies.

Categories and Subject Descriptors

I.4 [**Image Processing and Computer Vision**]: [Applications]; J.4 [**Social and Behavioral Sciences**]: [Psychology]

General Terms

Experimentation, Human Factors

Keywords

Personality modeling, BFI, Five-factor, Selfie, Images

ASM'15, October 30, 2015, Brisbane, Australia.
 ISBN 978-1-4503-3750-2/15/10 ...$15.00.
DOI: http://dx.doi.org/10.1145/2813524.2813528.

1. INTRODUCTION

'Selfie' refers to a self-potrait taken by a person. Selfies have become so popular that 2014 was dubbed as 'the year of the selfie' [1]. Social networking sites saw millions of selfies being uploaded (beginning from famous celebrities at important events like the Academy Awards, astronauts in space to anyone who has a smart-phone) [31] . Emergence of selfies as a new medium for self-expression and self-representation opens up the possibility of systematically assessing users' personality. Personality is defined as a series of "internal properties" that relate to overt behaviours [21]. Though there are many different theories which examine the predictive utility of personality, the Five Factor Model (FFM) [13] is one of the most widely used models. This consists of five dimensions namely openness to experience, conscientiousness, extroversion, agreeableness and neuroticism.

There are two dimensions to automatic personality prediction [30]: a) *personality recognition*, which deals with predicting personality of users which they externalize through distal cues, i.e., any form of observable behavior, and b) *personality perception*, which deals with what others perceive about the user through proximal cues. While automatic personality recognition is about inferring self-assessed personalities from machine detectable distal cues, automatic personality perception is about inferring the personality that observers attribute to a given user from proximal cues. A comprehensive review of these concepts is presented by [30].

A key factor to address the problem of automatic personality prediction, be it recognition or perception, is the source used for acquiring users' personality, i.e., the behavior we base the assessment upon. Studies in Psychology show that to get strong personality cues, users should be given the necessary freedom of control and motivation to express themselves through that behavior [14]. Many experimental studies have shown that standardized photos contain valid personality-related cues [7, 24, 25]. Similar work was conducted on profile photos and photos posted on social networks which provided evidence of personality expression in photos [16, 19].

However, when compared to other types of photos, selfies give individuals more freedom of controlling their face

Figure 1: **Selfies are a medium of self-expression:** Users can control what is in the selfie. Models can be trained to predict mid-level cues (See Table 1) in images which can predict users' personality. Cartoon source: [2]

visibility, emotional expression, and camera position, and are often posted on social media platforms used for self-presentation [22, 27].

Social Psychology research shows that there are several picture-coding cues (or mid-level cues) that are appropriate for coding selfies [16, 19, 25, 32, 28]. We consider some of them as listed in Table 1 and as shown in Figure 1. Using these, we attempt to computationally model users' personality based on their selfies. In particular, with selfies as the basis, we examine:

- if computer models can predict users' self-assessed personality (personality recognition)
- if computer models can predict how others assess users' personality (personality perception)
- what cues aid in both personality recognition and perception

Towards the above-mentioned goals, we build a selfie image dataset of 123 users from a popular microblogging website (Sina Weibo). The users also provide their self-assessed personality scores (using BFI-44 questionnaire [17]). We recruited eight research assistants from the Department of Psychology to provide their assessment of selfie subjects' personality using BFI-44. We also recruited two volunteers to annotate these images with mid-level cues.

Contributions: 1) We train models to detect mid-level cues (which are relevant to personality prediction based on psychology literature [16, 25]) in selfies, using low-level visual features. 2) The trained mid-level cue detectors are used to automatically predict users' personality, outperforming state-of-the-art features for most traits. 3) We present several insights on which mid-level cues contribute to personality recognition and personality perception.

2. RELATED WORK

Personality of users can be quantitatively measured and is found to be consistent w.r.t users behavior [21].

Table 1: Mid-Level Cues detected from Selfies

Mid-Level Cue	Label Description
Face Visibility	Not Visible/Partial/Complete
Photoshop Editing	Present/Absent
Public Location	Yes/No
Private Location	Yes/No
Duckface	Present/Absent
Pressed Lips	Present/Absent
Emotional Positivity	Negative/Neutral/Positive
Alone	Yes/No
Amount of Body	Face Only/Shoulder-up/Waist-up
Eyes Looking at Cam.	Yes/No
Camera Height	Below/At/Above Head-level
Camera In Front	Yes/No
Gender	Male/Female
Age	Below 18/18-20/21-25/Above 26

According to Brunswick Lens model [9], personality is externalized through 'distal cues' (i.e., any behavior that others can perceive), enabling others to form impression about the users' personality. Researchers used cues like choice of words in language [5], intonation of voice while speaking [23], kind of photos one likes [15], type of people one befriends [12], other forms of multimodal information [4, 10, 29] etc., to model users' personality (see [30] for a thorough review). With images becoming a wide-spread channel for communication and expression, there has been a recent trend to use facial images from Facebook [11], random portraits on the web [26] and existing face recognition datasets [3] to model users' personality.

However, face recognition datasets [3] and random portraits on the web [26] would not have provided subjects in the images an appropriate platform for self-expression and freedom of control [16], which are necessary cues for personality to be strongly expressed [14].

Therefore we use selfies of users and propose a novel approach to build mid-level cue detectors (which are especially relevant for selfies and facial images [16, 25])

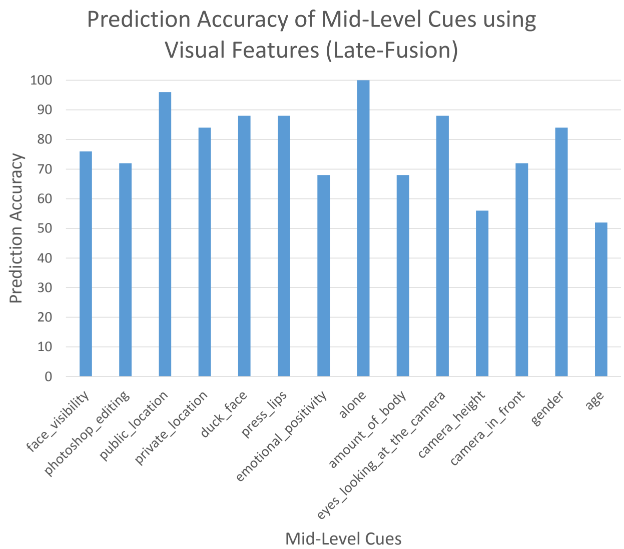

Figure 2: **Prediction Accuracy of Mid-Level cues:** Low-level visual features (listed in Section 4) are used to train models which detect mid-level cues in selfies. All models are combined (decision-fusion) for final prediction.

using visual features. These detectors are then used to model users' personality. To the best of our knowledge, no previous work has targeted the task of building such mid-level cue detectors for facial images (though similar approaches are shown to be successful in domains such as sentiment analysis etc. [8]).

3. DATA COLLECTION

To participate in our study, we recruited 612 users from Sina Weibo (a popular microblogging website similar to twitter in China). Previous studies show that profile photos may have more impact on personality judgements than textual self-disclosures [11]. Therefore, profile pictures of all participants were downloaded. Two independent raters identified 123 pictures (out of 612) to be selfie portraits, which are used in our study to model personality.

Each of the 123 users (89 females and 34 males) answered the BFI-44 [17] and gave us their profile user names, usage frequency and other basic information (namely gender, age, country of residence and ethnicity). The scores on the five factors were calculated using BFI scheme (ranging between 1 and 5).

In this work, we aim to build mid-level detectors to automatically predict both self-assessed and others-assessed personality scores of users. To get the ground truth on others-assessed personality of users, eight research assistants from the Department of Psychology viewed each user's selfie and rated their impression of the selfie subjects personality using the same BFI questionnaire (See [30] for more details on formats). To get the ground truth for learning the mid-level cue detectors, two independent raters annotated the selfies with the mid-level cues listed in Table 1. The two raters achieved a 90.81% inter-rater agreeability ($p<0.001$) in their coding.

4. LEARNING MID-LEVEL CUE DETECTORS

4.1 Feature Extraction

We extracted 10 different visual features motivated by their relevance to capturing different aspects of the images' characteristics.

Color Histograms: in RGB space is evaluated, as color is an important clue in conveying one's preferences. These are supposed to represent users' inclination to different colors.

Aesthetic Features [20]: are used to characterise photographic styles (e.g. rule-of-thirds, vanishing points, etc.) based on art theory and psychological studies. These are supposed to give an idea about users' aesthetic preferences.

GIST: Spectral information and course localization are used to represent global structure of the scene in an image. These convey information about the scene in the selfies.

LBP: LBP is used to encode visual texture perception information. As most of the area in selfies is occupied by users' faces, LBP features represent facial information as well.

Attribute Features [34]: are useful for characterizing abstract sentiments associated with the image by automatically designing attributes encoded by a compact category-attribute matrix.

Bag-of-Visual-Words [33]: A vocabulary is generated by vector quantization of keypoint descriptors, which in this case is taken as 1500 dimension.

Fisher encodings of SIFT, SURF, HOG and MSER: The Fisher Vector encoding extends the Bag-of-Visual-Words by going beyond 0-order statistics and by encoding the second order statistics about each visual word's local descriptors' distribution.

4.2 Training Mid-Level Cue Detectors

After extracting the features from images, we used LibSVM with RBF kernel to train mid-level cue detectors. Parameter tuning was performed on a 5-fold cross validation and results are reported on 25 % test data split. We use the output of mid-level detectors as features to predict the scores on personality (both self-assessed and others-assessed) and we compare the results with those achieved using visual features. Since personality scores were continuous values between 1 and 5, we employed LASSO formulation to model them.

5. RESULTS & DISCUSSION

5.1 Mid-Level Cue Detectors for Selfies

The accuracy of detecting mid-level cues using visual features (late-fusion) is shown in Figure 2. Each of the visual features were used to train models to detect mid-level cues and then all the models were combined using a late-fusion technique. Cues such as alone and public_location are detected with >75% accuracy with almost all the features. However, other cues need specific features for good detection performance. For example pressed lips is detected the best by LBP (followed by BOW), which is known to capture facial features well. Similarly emotional positivity is best detected by

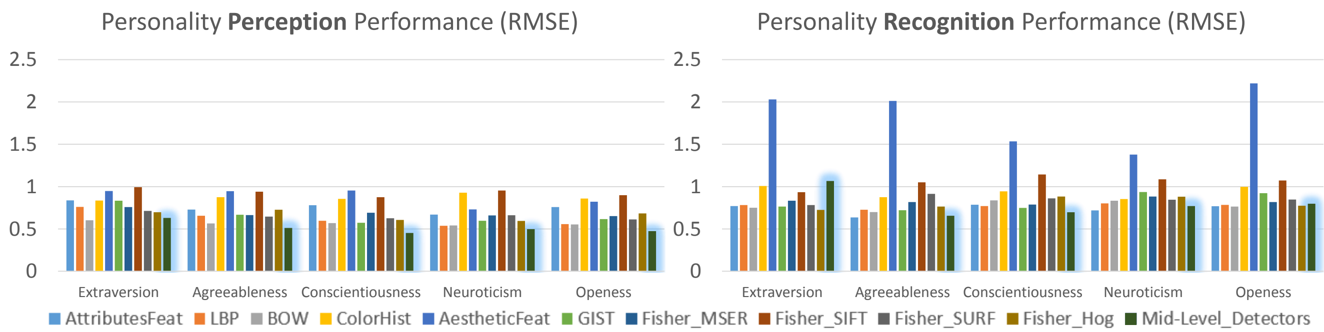

Figure 3: Performance (RMSE) of various features at predicting personality scores. Mid-level detectors outperform other features at predicting most traits

Fisher_SIFT and Aesthetic Features (which are designed specifically for emotion detection in natural images).

Age and camera height are the least performing detectors, probably because of the complex nature of age estimation and the absence of specific features to detect pose respectively. The final mid-level cue detectors we use for personality modeling are built using a late-fusion scheme of all features. It should be noted that mid-level detectors are not cent percent accurate (that is there exists a confidence level to their predictions) and consequent errors would propagate to personality predictions as well.

5.2 Predicting Personality

The results of predicting users' personality using mid-level cue detectors are shown in Figure 3. In predicting self-assessed personality (recognition) we find that mid-level cue detectors outperformed other features in agreeableness, conscientiousness and neuroticism. And in predicting others-assessed personality (perception), they outperformed other features in all traits except extraversion. In both case, they outperform the features LBP and SIFT, HOG (Fisher encodings), which are used by previous work [11, 26] on predicting personality based on facial images.

The lower performance in personality recognition, when compared to personality perception, might be due to users' manipulation of their self-presentation to create a socially desirable self-image [22]. This implies that there is a possibility of users' exaggerating their expressions (and other mid-level cues), thereby leading to higher errors in personality recognition. However, this exaggeration might be visible to a person who is perceiving the users' personality, leading to a stronger association between mid-level cues and personality perception. There also exists the biases present in selfies posted on social media - most of the users seem to be smiling and having a positive look, thereby reducing the diversity of the self-presentations. This might be another reason for performance on personality perception to be higher.

To further understand which specific mid-level cue detectors were significant in predicting self-assessed traits, we performed correlation analysis (shown in Tables 2). The following are some observations:

1. Contradictory to a previous psychology finding where extraversion was related to positive emotional expression [24], it was not significantly correlated with any cue in our study. One reason could be that users, irrespective of their degree of extraversion, tend to show positive emotion in selfies to maintain positive self-impression [22].
2. Agreeableness was negatively correlated with camera height indicating that agreeable individuals more likely take selfies from below.
3. Conscientiousness was negatively correlated with private location indicating that conscientious people do not like to expose their private space in the

Table 2: Significant Correlations ($p < 0.01$) b/w Self-Assessed Personality Traits and Mid-Level Cues

Self-Assessed trait	Sig. Cue	Corr.
Extraversion	-	-
Agreeableness	Emotional Positivity	0.18
	Camera Height	-0.20
Conscientiousness	Private Location	-0.20
Neuroticism	Duck Face	0.21
Openness	Emotional Positivity	0.22

Table 3: Significant Correlations ($p < 0.01$) b/w Others-Assessed Personality Traits and Mid-Level Cues

Others-Assessed trait	Sig. Cue	Corr.
Extraversion	Pressed Lips	-0.19
	Emotional Positivity	0.29
Agreeableness	Emotional Positivity	0.50
	Eyes Look at Camera	0.24
Conscientiousness	Duck Face	-0.31
	Emotional Positivity	0.25
	Location Information	0.30
	Public Location	0.25
	Photoshop Editing	-0.20
Neuroticism	Duck Face	0.25
	Emotional Positivity	-0.40
	Face Visibility	-0.21
	Amount of Body	-0.22
	Alone	0.22
	Location Information	-0.19
Openness	Emotional Positivity	-0.22
	Eyes Look at Camera	0.21
	Face Visibility	-0.26

background as corroborated by a psychology finding [18].
4. Neuroticism positively correlated with duckface indicating.
5. Openness positively correlated with emotional positivity, which was not observed in previous findings.

A similar analysis on others-assessed traits, to identify cues that observers used when judging personality, revealed the following:

1. Rating on extraversion was negatively correlated with pressed lips, possibly because it's being seen as a symptom of shyness. Extraversion was positively correlated with emotional positivity, supporitng the relationship between perception of extraversion and smiling [24].
2. Agreeableness ratings were positively correlated with eyes looking at camera, supporting that users who had eye contact are seen as more agreeable [6].
3. Rating on neuroticism was positively correlated with duckface and negatively correlated with face visibility, indicating that showing a full duckface implies that the user is moody.

6. CONCLUSION

In this paper, we built mid-level cue detectors for selfie portrait images by exploiting several visual features. These mid-level detectors were then used to model personality of users in the selfies. We also performed correlation analysis on the different mid-level cues and the personality traits to understand selfie-taking behaviors of users with different personalities and how others would perceive the users based on their selfies.

It would be interesting to emperically verify how reliable selfies are at personality prediction and how significant is the degree of exaggeration (and its affect on personality recognition), when compared to other types of photos of the users. We are examining these aspects in our ongoing work. It would also be interesting to explore the degree of exaggeration that users exhibit w.r.t different mid-level cues (for example, if the user is posing in a duckface to an extent that it looks awkward and uncomfortable vs. otherwise).

7. REFERENCES

[1] BBC News blog. *http://www.bbc.com/news/blogs-trending-30394875*.

[2] Pinterest. *https://www.pinterest.com/LokidByLoki/angry/*.

[3] N. Al Moubayed, Y. Vazquez-Alvarez, A. McKay, and A. Vinciarelli. Face-based automatic personality perception. In *ACM Multimedia*, 2014.

[4] F. Alam and G. Riccardi. Predicting personality traits using multimodal information. In *Proceedings of the 2014 ACM Multi Media on Workshop on Computational Personality Recognition*, pages 15–18. ACM, 2014.

[5] S. Argamon, S. Dhawle, M. Koppel, and J. Pennebaker. Lexical predictors of personality type. 2005.

[6] J. E. Bono and T. A. Judge. Personality and transformational and transactional leadership: a meta-analysis. *Journal of applied psychology*, 89(5):901, 2004.

[7] P. Borkenau, S. Brecke, C. Möttig, and M. Paelecke. Extraversion is accurately perceived after a 50-ms exposure to a face. *Journal of Research in Personality*, 43(4):703–706, 2009.

[8] D. Borth and et. al. Sentibank: large-scale ontology and classifiers for detecting sentiment and emotions in visual content. In *ACM Multimedia*, 2013.

[9] E. Brunswik. *Perception and the representative design of psychological experiments*. U. of Cal. Press, 1956.

[10] O. Celiktutan, F. Eyben, E. Sariyanidi, H. Gunes, and B. Schuller. Maptraits 2014-the first audio/visual mapping personality traits challenge-an introduction: Perceived personality and social dimensions. In *Proceedings of the 16th International Conference on Multimodal Interaction*, pages 529–530. ACM, 2014.

[11] F. Celli, E. Bruni, and B. Lepri. Automatic Personality and Interaction Style Recognition from Facebook Profile Pictures. In *ACM Multimedia*. ACM, 2014.

[12] J. Golbeck, C. Robles, and K. Turner. Predicting personality with social media. In *CHI'11 extended abstracts*. ACM, 2011.

[13] L. R. Goldberg. An alternative" description of personality": the big-five factor structure. *Journal of personality and social psychology*, 59(6):1216, 1990.

[14] S. D. Gosling, S. J. Ko, T. Mannarelli, and M. E. Morris. A room with a cue: personality judgments based on offices and bedrooms. *Journal of personality and social psychology*, 82(3):379, 2002.

[15] S. C. Guntuku, S. Roy, and W. Lin. Personality modeling based image recommendation. In *MultiMedia Modeling*, pages 171–182. Springer, 2015.

[16] J. A. Hall and N. Pennington. Self-monitoring, honesty, and cue use on facebook: The relationship with user extraversion and conscientiousness. *Computers in Human Behavior*, 2013.

[17] O. P. John and S. Srivastava. The big five trait taxonomy: History, measurement, and theoretical perspectives. *Handbook of personality: Theory and research*, 2(1999):102–138, 1999.

[18] I. A. Junglas, N. A. Johnson, and C. Spitzmüller. Personality traits and concern for privacy: an empirical study in the context of location-based services. *European Journal of Information Systems*, 17(4):387–402, 2008.

[19] N. C. Krämer and S. Winter. Impression management 2.0. *Journal of Media Psychology: Theories, Methods, and Applications*, 20(3):106–116, 2008.

[20] J. Machajdik and A. Hanbury. Affective image classification using features inspired by psychology and art theory. In *ACM Multimedia*, 2010.

[21] G. Matthews, I. J. Deary, and M. C. Whiteman. *Personality traits*. Cambridge University Press, 2003.

[22] S. Mehdizadeh. Self-presentation 2.0: Narcissism and self-esteem on facebook. *Cyberpsychology, Behavior, and Social Networking*, 13(4):357–364, 2010.

[23] G. Mohammadi and A. Vinciarelli. Humans as feature extractors: combining prosody and personality

perception for improved speaking style recognition. In *IEEE SMC*, 2011.

[24] L. P. Naumann, S. Vazire, P. J. Rentfrow, and S. D. Gosling. Personality judgments based on physical appearance. *Personality and Social Psychology Bulletin*, 2009.

[25] S. Nestler, B. Egloff, A. C. Küfner, and M. D. Back. An integrative lens model approach to bias and accuracy in human inferences: hindsight effects and knowledge updating in personality judgments. *Journal of personality and social psychology*, 103(4):689, 2012.

[26] J. Nie, P. Cui, Y. Yan, L. Huang, Z. Li, and Z. Wei. How your portrait impresses people?: Inferring personality impressions from portrait contents. In *ACM Multimedia*, 2014.

[27] Z. Papacharissi. A networked self. *A Networked Self*, page 304, 2011.

[28] L. Qiu, J. Lu, S. Yang, W. Qu, and T. Zhu. What does your selfie say about you? *Computers in Human Behavior*, 52:443–449, 2015.

[29] R. Srivastava and et. al. Don't ask me what i'm like, just watch and listen. In *ACM Multimedia*, 2012.

[30] A. Vinciarelli and G. Mohammadi. A survey of personality computing. 2014.

[31] S. Vranica. Behind the Preplanned Oscar Selfie: Samsung's Ad Strategy. *Wall Street Journal*, 6:41, 2014.

[32] Y.-H. Wang. *Taiwanese girls' self-portraiture on a social networking site*. PhD thesis, London School of Economics, 2012.

[33] J. Yang, Y.-G. Jiang, A. G. Hauptmann, and C.-W. Ngo. Evaluating bag-of-visual-words representations in scene classification. In *Proceedings of the international workshop on Workshop on multimedia information retrieval*, pages 197–206. ACM, 2007.

[34] F. X. Yu and et.al. Designing category-level attributes for discriminative visual recognition. In *CVPR*. IEEE, 2013.