# Aesthetic Evaluation and Guidance for Mobile Photography

Hao Lou, Heng Huang*
Chaoen Xiao and Xin Jin†
Beijing Electronic Science and Technology Institute
Fengtai District, Beijing 100070, China


**Figure 1: Our proposed Aesthetic Dashboard: in the left-up conner, the overall aesthetic score and the aesthetic attribute radar map (light, composition, color) are shown. Using this sub-dashboard, the novices can simply adjust their shooting according to the scores, the higher the better. In the right part, more complicated aesthetic attribute dashboards are shown. The amateurs or professional users can refer the template matching scores to obtain the desired patterns of light, color and composition. For portraits we also show the guidance of face light, body pose and the garment color, see the supplementary material.**

## ABSTRACT

Nowadays, almost everyone can shoot photos using smart phones. However, not everyone can take good photos. We propose to use computational aesthetics to automatically teach people without photography training to take excellent photos. We present *Aesthetic Dashboard*: a system of rich aesthetic evaluation and guidance for mobile photography. We take 2 most used types of photos: landscapes and portraits into consideration. When people take photos in the preview mode, for landscapes, we show the overall aesthetic score and scores of 3 basic attributes: light, composition and color usage. Meanwhile, the matching scores of the 3 basic attributes of current preview to typical templates are shown, which can help users to adjust 3 basic attributes accordingly. For portraits, besides the above basic attributes, the facial appearance, the guidance of face light, body pose and the garment color are also shown to the users. This is the first system that can teach mobile users to shoot good photos in the form of aesthetic dashboard, through which, users can adjust several aesthetic attributes to take good photos easily.

## CCS CONCEPTS

• **Applied computing → Media arts**.

## KEYWORDS

image aesthetics, mobile photography, aesthetic evaluation, aesthetic guidance, aesthetic dashboard

**ACM Reference Format:**

Hao Lou, Heng Huang and Chaoen Xiao and Xin Jin. 2021. Aesthetic Evaluation and Guidance for Mobile Photography. In *Proceedings of the 29th ACM International Conference on Multimedia (MM '21), October 20–24, 2021, Virtual Event, China.* ACM, New York, NY, USA, 3 pages. https://doi.org/10.1145/3474085.3478557

*Both authors contributed equally to this research.
†Corresponding author email: jinxin@besti.edu.cn.


*MM '21, October 20–24, 2021, Virtual Event, China*

ACM ISBN 978-1-4503-8651-7/21/10.
https://doi.org/10.1145/3474085.3478557

## 1 INTRODUCTION

Taking a photo is quite convenient by using smart phones. Trillions of photos are taken everyday by our smart phones. The resolution of these photos are getting higher. However, most of users have not been trained on photography. They do not know how to shot a good photo by choosing appropriate composition, light or color.

Recently, several assistance systems [7, 13–15] for mobile photography and aesthetic assessment methods [4, 5, 9] are proposed. Rawat et al. [13] use photos from social media to suggest the best shooting place and angle for mobile photography. The shot suggestion function of Samsung Galaxy S10 [14] shows the guidance line and circle to guide users to obtain good composition. Google proposes camera view adjustment prediction [15] for improving photo composition for mobile photography. The PicMe system proposed by Kim et al. [7] use interactive visual guidance for taking requested photo composition by another user.

Most of the related shooting assistance systems only consider simple composition guidance. None of them explicitly show the aesthetic scores of both the overall and the aesthetic attributes. We propose the aesthetic dashboard which shows the rich aesthetic evaluation and guidance for the mobile users to shoot excellent photos. Using our aesthetic scores, the novices can simply adjust their shooting according to the scores, the higher the better. The amateurs or professional users can refer the template matching scores to obtain the desired patterns of light, color and composition. For portraits, we also show facial appearance and the guidance of face light, body pose and the garment color.

## 2 AESTHETIC DASHBOARD

### 2.1 Landscapes

*2.1.1 Overall Aesthetic Score.* We use a multi-task network structure with Efficientnet [16] as the backdone network to extract aesthetic depth features of landscape images, We extract aesthetic abstract features through aesthetic total score ten-class training. Then we fuse the handcraft features of various attributes in the multi-task network and obtain the aesthetic total-score of the image according to the regression training.

*2.1.2 Attributes Score.* In terms of attributes scoring, we obtain aesthetic abstract feature maps from total score training. Combining the channel attention [18], we extract abstract features for composition, light and color. In addition, we further process these features and each finally contains feature of 10 dimensions. In the process of attribute regression, we add a small number of hand-craft features to the final full connection layers, Finally, we incorporate these hand-craft features into the 10 high-dimensional attribute features for the regression of attribute scores, The results are presented in the form of the radar graph.

*2.1.3 Composition Templates.* We summarize ten common composition templates. They are: gold section, central composition, slkew composition, triangle composition, guide line, rule of thirds, symmetrical composition, Diagomal Composition, frame composition and round composition, For each landscape image, we preprocess it through the silent area detection [12] and edge extraction [12], then we obtain the composition information according to the center location of the silent area, the primal sketch [3] and circle area.

*2.1.4 Light Templates.* We use the solar-sky model [10] to display outdoor illumination distribution information. We divided the sky into 32 different regions in the range of level 360 degrees and obtain the solar probability for the forward light, side light, back light and side back light.

*2.1.5 Color Templates.* We use the color harmony model [2] to display the color distribution. We divide the histogram according to the hue channel in HSV, and calculate the different harmonious color models in type i, V, L, I, T, Y, X and N of color harmony. Finally, we calculate the average and maximum of hue in different templates to evaluate the color matching degree.

### 2.2 Portraits

*2.2.1 Overall Aesthetic Score.* We got the overall aesthetic score by integrating four attributes including lighting, composition, color and appearance.

*2.2.2 Composition Score.* We calculated seven composition features of the picture, including trisection, golden section, diagonal, central composition, triangle composition, L-shape composition and symmetrical composition. The composition score is calculated according to the correlation between the image and the composition template. We used the Efficient-Det [17] to get the portrait rectangle.

*2.2.3 Light Score.* We divide the image into 64 pieces and the portrait into 8 pieces to calculate the lighting features. And the features are used as input of neural network to train in the data set of AADB [8], PCCD [1], EVA [6], PADB to get the light score.

*2.2.4 Color Score.* We divide the image into 64 blocks and count the color features. Then we use neural network to train on the attribute data set to get the color score.

*2.2.5 Composition Templates.* The composition templates used in this demo include trisection, golden section, diagonal, center composition, triangle composition, L-shape composition and symmetrical composition.

*2.2.6 Light Templates.* We used DPR [20] to get the radiance map we need. The lighting templates used in this demo include butterfly light, side light and Rembrandt light.

*2.2.7 Pose Templates.* This demo collects a number of commonly used pose as pose templates to guide shooting. And we used the detectron2 [19] to get the skeleton.

*2.2.8 Appearance.* This demo uses the neural network model trained on the SCUT-FBP5500 [11]

## ACKNOWLEDGEMENTS

We thank the ACs, reviewers, and the advices from the collaborators of the Huawei and Shaoding.ai companies. This work is partially supported by the National Natural Science Foundation of China (62072014), the Beijing Natural Science Foundation (L192040), and the Advanced Discipline Construction Project of Beijing Universities (20210041Z0401).

## REFERENCES


[1] Kuang-Yu Chang, Kung-Hung Lu, and Chu-Song Chen. 2017. Aesthetic Critiques Generation for Photos. In *IEEE International Conference on Computer Vision, ICCV 2017, Venice, Italy, October 22-29, 2017*. IEEE Computer Society, 3534–3543. https://doi.org/10.1109/ICCV.2017.380

[2] Daniel Cohen-Or, Olga Sorkine, Ran Gal, Tommer Leyvand, and Ying-Qing Xu. 2006. Color harmonization. *ACM Trans. Graph.* 25, 3 (2006), 624–630. https://doi.org/10.1145/1141911.1141933

[3] Qi Han, Kai Zhao, Jun Xu, and Ming-Ming Cheng. 2020. Deep Hough Transform for Semantic Line Detection. In *Computer Vision - ECCV 2020 - 16th European Conference, Glasgow, UK, August 23-28, 2020, Proceedings, Part IX (Lecture Notes in Computer Science, Vol. 12354)*, Andrea Vedaldi, Horst Bischof, Thomas Brox, and Jan-Michael Frahm (Eds.). Springer, 249–265. https://doi.org/10.1007/978-3-030-58545-7_15

[4] Xin Jin, Le Wu, Xiaodong Li, Siyu Chen, Siwei Peng, Jingying Chi, Shiming Ge, Chenggen Song, and Geng Zhao. 2018. Predicting Aesthetic Score Distribution Through Cumulative Jensen-Shannon Divergence. In *Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence, New Orleans, Louisiana, USA, February 2-7, 2018*. https://www.aaai.org/ocs/index.php/AAAI/AAAI18/paper/view/16074

[5] Xin Jin, Le Wu, Geng Zhao, Xiaodong Li, Xiaokun Zhang, Shiming Ge, Dongqing Zou, Bin Zhou, and Xinghui Zhou. 2019. Aesthetic Attributes Assessment of Images. In *Proceedings of the 27th ACM International Conference on Multimedia, MM 2019, Nice, France, October 21-25, 2019*, Laurent Amsaleg, Benoit Huet, Martha A. Larson, Guillaume Gravier, Hayley Hung, Chong-Wah Ngo, and Wei Tsang Ooi (Eds.). ACM, 311–319. https://doi.org/10.1145/3343031.3350970

[6] Chen Kang, Giuseppe Valenzise, and Frédéric Dufaux. 2020. EVA: An Explainable Visual Aesthetics Dataset. In *Joint Workshop on Aesthetic and Technical Quality Assessment of Multimedia and Media Analytics for Societal Trends* (Seattle, WA, USA) *(ATQAM/MAST'20)*. Association for Computing Machinery, New York, NY, USA, 5–13. https://doi.org/10.1145/3423268.3423590

[7] Minju Kim and Jungjin Lee. 2019. PicMe: Interactive Visual Guidance for Taking Requested Photo Composition. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, CHI 2019, Glasgow, Scotland, UK, May 04-09, 2019*, Stephen A. Brewster, Geraldine Fitzpatrick, Anna L. Cox, and Vassilis Kostakos (Eds.). ACM, 395. https://doi.org/10.1145/3290605.3300625

[8] Shu Kong, Xiaohui Shen, Zhe Lin, Radomir Mech, and Charless Fowlkes. 2016. Photo Aesthetics Ranking Network with Attributes and Content Adaptation. In *ECCV*.

[9] Qi Kuang, Xin Jin, Qinping Zhao, and Bin Zhou. 2020. Deep Multimodality Learning for UAV Video Aesthetic Quality Assessment. *IEEE Trans. Multim.* 22, 10 (2020), 2623–2634. https://doi.org/10.1109/TMM.2019.2960656

[10] Jean-François Lalonde, Alexei A. Efros, and Srinivasa G. Narasimhan. 2012. Estimating the Natural Illumination Conditions from a Single Outdoor Image. *Int. J. Comput. Vis.* 98, 2 (2012), 123–145. https://doi.org/10.1007/s11263-011-0501-8

[11] L. Liang, L. Lin, L. Jin, D. Xie, and M. Li. 2018. SCUT-FBP5500: A Diverse Benchmark Dataset for Multi-Paradigm Facial Beauty Prediction. In *2018 24th International Conference on Pattern Recognition (ICPR)*.

[12] Jiangjiang Liu, Qibin Hou, and Ming-Ming Cheng. 2020. Dynamic Feature Integration for Simultaneous Detection of Salient Object, Edge, and Skeleton. *IEEE Trans. Image Process.* 29 (2020), 8652–8667. https://doi.org/10.1109/TIP.2020.3017352

[13] Yogesh Singh Rawat and Mohan S. Kankanhalli. 2017. ClickSmart: A Context-Aware Viewpoint Recommendation System for Mobile Photography. *IEEE Trans. Circuits Syst. Video Technol.* 27, 1 (2017), 149–158. https://doi.org/10.1109/TCSVT.2016.2555658

[14] Samsung. [n.d.]. How to use Shot suggestions on the Galaxy S10. [EB/OL]. https://www.samsung.com/levant/support/mobile-devices/how-to-use-shot-suggestions/ Accessed September 22, 2020.

[15] Yu-Chuan Su, Raviteja Vemulapalli, Ben Weiss, Chun-Te Chu, Philip Andrew Mansfield, Lior Shapira, and Colvin Pitts. 2021. Camera View Adjustment Prediction for Improving Image Composition. *CoRR* abs/2104.07608 (2021). arXiv:2104.07608 https://arxiv.org/abs/2104.07608

[16] Mingxing Tan and Quoc V. Le. 2019. EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks. In *Proceedings of the 36th International Conference on Machine Learning, ICML 2019, 9-15 June 2019, Long Beach, California, USA (Proceedings of Machine Learning Research, Vol. 97)*, Kamalika Chaudhuri and Ruslan Salakhutdinov (Eds.). PMLR, 6105–6114. http://proceedings.mlr.press/v97/tan19a.html

[17] Mingxing Tan, Ruoming Pang, and Quoc V. Le. 2020. EfficientDet: Scalable and Efficient Object Detection. In *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition, CVPR 2020, Seattle, WA, USA, June 13-19, 2020*. IEEE, 10778–10787. https://doi.org/10.1109/CVPR42600.2020.01079

[18] Qilong Wang, Banggu Wu, Pengfei Zhu, Peihua Li, Wangmeng Zuo, and Qinghua Hu. 2020. ECA-Net: Efficient Channel Attention for Deep Convolutional Neural Networks. In *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition, CVPR 2020, Seattle, WA, USA, June 13-19, 2020*. IEEE, 11531–11539. https://doi.org/10.1109/CVPR42600.2020.01155

[19] Yuxin Wu, Alexander Kirillov, Francisco Massa, Wan-Yen Lo, and Ross Girshick. 2019. Detectron2. https://github.com/facebookresearch/detectron2.

[20] Hao Zhou, Sunil Hadap, Kalyan Sunkavalli, and David Jacobs. 2019. Deep Single-Image Portrait Relighting. In *2019 IEEE/CVF International Conference on Computer Vision, ICCV 2019, Seoul, Korea (South), October 27 - November 2, 2019*. IEEE, 7193–7201. https://doi.org/10.1109/ICCV.2019.00729