An Esthetics Rule-Based Ranking System for Amateur Photos

Che-Hua Yeh[†] Wai-Seng Ng[†] Brian A. Barsky* Ming Ouhyoung[†]
[†]National Taiwan University [*]University of California, Berkeley

1. Introduction

With the current widespread use of digital cameras, the process of selecting and maintaining personal photos is becoming an onerous task. To our knowledge, there has been little research on photo evaluation based on computational esthetics. Photographers around the world have established some general rules for taking good photos. Building upon artistic theories and human visual perception is difficult since the results tend to be subjective. Although automatically ranking award-wining professional photos may not be a sensible pursuit, such an approach may be reasonable for photos taken by amateurs. In the next section, we introduce rules for such a system.

2. Rules of Esthetics

Rules of esthetics in photography describe how to arrange different visual elements inside an image frame. We categorize these rules into two major parts: photo composition and color distribution.

2.1. Photo Composition

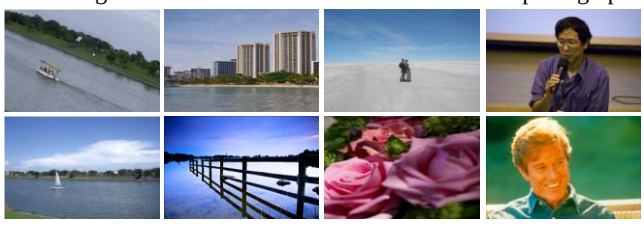
a. Horizontal balance. To obtain geometrical balance in a photo, horizon should be level. Fig. 1(a) shows unlevel and level horizons in the top and bottom photos, respectively.

b. Line patterns. Perspective viewing for the built environment can reveal parallel and radial line patterns, enabling observers to create a mental 3D model. Fig. 1(b) one shows the case of parallel and perspective lines horizons in the top and bottom photos, respectively. We extract parallel and perspective lines separately.

c. Size of ROIs. Simplicity is a distinguishing factor in determining the quality of a photo. We use the normalized size of the ROIs in a photo, since a user always looks at the ROI. Fig. 1(c) show different ROI sizes, too small and too big in the top and bottom photos, respectively.

d. Merger avoidance. A merger occurs when two or more separate objects overlap in a photo. This can cause important scene elements to lose their significance. In our work, **merger avoidance** attempts to avoid the case that background lines intersect a human head, as shown in the top of Fig 1(d), which is sometimes taboo in Asian cultures.

e. Selecting Location of Region of Interest (ROI) -- the rule of thirds. The rule of thirds states that the ROI should be placed one third along the horizontal or vertical dimension of the photograph.



(a) Horizontal balance (b) Line Patterns (c) Size of ROIs (d) Merger avoidance

Fig. 1 Rules of Photo Composition

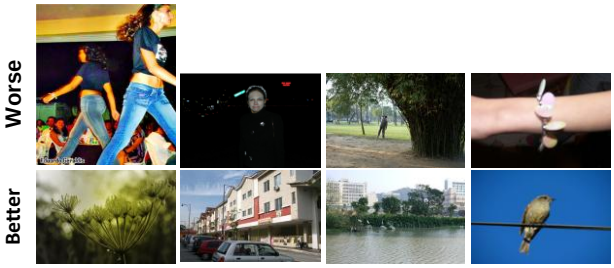
2.2. Color and Intensity Distribution

a. Color harmonization. The color harmonization technique [1] is used to measure the quality of color distribution. The top and bottom images of Fig. 2(a) show the worse and better photos, respectively.

b. Contrast. Contrast is a low-level feature which is useful for evaluating image quality. High contrast is usually better than low contrast, as can be seen in Fig. 2(b).

c. Intensity balance. Visual elements within a photo can provide observers with a sense of weight. Balancing the total weight of a photo is important, as shown in Fig. 2(c).

d. Blurriness. The ROI in an image should generally be in sharp focus, as shown in Fig. 2(d).



(a) Color Harmonization (b) Contrast (c) Intensity balance (d) Blurriness

Fig. 2 Rules of Color and Intensity Distribution

3. Automatic Ranking of Photos

We validated our approach automatically using support vector machine (SVM) training and testing as well as by using human subjects in a user study.

To perform automated verification, 1000 top and 1000 bottom scored photos were collected from ("dpchallenge") http://www.dpchallenge.com/. Half the photos were selected for support vector machine (SVM) training data, and then the others were used for testing. This was performed using LIBSVM. Ten features, based on the rules described above, were extracted from each photo. After the SVM training process with cross validation, our final output accuracy was 81%, which improved upon the 72% to 76% reported by Ke et al. at similar recall rate [2].

A user study was also performed to validate our system with human subjects. First, 10,000 photos were randomly collected from the Flickr website. Based on our system's ranking of these photos, the 50 highest and 50 lowest ranked photos are then selected for the user study. The total execution time of the Support Vector Regression analysis for predicting 10,000 photos, each with 10-D feature vectors, was 3.5 seconds, using an Intel CoreQuad 2.4 GHz PC. At an average 0.35 milliseconds per photo, our system is quite efficient.

In the user study, users are provided with pairs of photos, where one was from the highest ranked group and the other was from the lowest ranked group, and they are asked which of the two photos they prefer. Using the F-test to analyze the data of user study corroborated our hypothesis at 99% threshold that our system can effectively rank photos taken by amateurs.

4. References

[1] Daniel Cohen-Or, Olga Sorkine, Ran Gal, Tommer Leyvand, and Ying-Qing Xu. "Color Harmonization", *SIGGRAPH 2006 Conference Proceedings*, pp. 624-630.

[2] Yan Ke, Xiaoou Tang, and Feng Jing. "The Design of High-Level Features for Photo Quality Assessment", *Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, New York, 17-22 June 2006, IEEE Computer Society, pp. 419-426.

SIGGRAPH Asia 2009, Yokohama, Japan, December 16–19, 2009.
ISBN 978-1-60558-858-2/09/0012