Evaluating photo aesthetics using machine learning

Domen Pogačnik, Robert Ravnik, Narvika Bovcon, Franc Solina
Faculty of Computer and Information Science, University of Ljubljana
domen.pogacnik@gmail.com,
{robert.ravnik, narvika.bovcon, franc.solina}@fri.uni-lj.si

ABSTRACT

In this paper we propose a method for automatic assessment of aesthetic appeal of photographs. We identify significant parameters that distinguish high quality photography from low quality snapshots. On the basis of these parameters, we defined calculable features for automatic assessment of photography aesthetics using machine learning methods. The calculation of features depends heavily on the identification of the subject in photographs. With the subject identified, we defined and implemented various features to analyze various aspects of a photograph. The features were tested on two datasets. First dataset was obtained from Flickr and manually labeled for evaluation. Second dataset was based on photographs from DPChallenge portal where subjects were identified with a face detection algorithm. Both experiments showed some promising results. In this article we specify the features which contribute to a successful classification of photographs, analyze their influence and discuss the results. In conclusion, we offer some suggestions for further research.

1 INTRODUCTION

With ever-decreasing cost of digital photography along with ease of their manipulation, the volume of photographs taken has increased exponentially. According to statistics quoted by the largest social network – Facebook, an average of 200 million photographs are uploaded daily by its users. Evidently, there is a need for an automatic system that rates and manages such content. Therefore, photo-quality evaluation is an area that has particularly attracted recent research attention.

Evaluating photo aesthetics proved to be challenging because of several reasons. The notion of a *high quality* photograph as perceived by a viewer is often an abstract concept. Even experienced photographers can disagree on the quality of a particular photo. However, there are some common rules that most of high quality photos follow. For instance, photos taken by experienced photographers usually adhere to one of the composition rules and color selection, which makes them more visually appealing than those taken by amateurs. Therefore, obtaining reliable evaluations of photo quality was an important motivation for our work.

We formalized abstract concepts of photographic quality into quality measures and designed features that can be efficiently determined by computers. Our method is also a top-down approach such as the one by Ke et al. [1] where high-level semantic features for photo quality assessment are constructed. Each constructed feature correlates to a concept established in professional photography.

The article is organized as follows: related work is presented in Sect. 2, parameters of high quality photographs in Sect. 3, followed by the proposed features in Sect. 4. Evaluation datasets and experimenal results are presented in Sect. 5 and 6, respectively. Sect. 7 concludes the paper with discussion and future work.

2 RELATED WORK

In recent research of automatic evaluation of photography aesthetics, Ke et al. [1] employed a top-down approach to construct high-level features for photo quality evaluation. Features are close to the concepts described by experienced photographers and were extracted from low level cues like noise, blur, color, brightness, contrast and spatial distribution of edges. In addition to the low level cues, the significance of complex features such as image similarity, region composition and depth of field indicator are presented in [2, 3]. Boutell and Luo [4] studied the impact of metadata recorded by a camera at image conception. The metadata consisted of camera settings such as ISO speed rating, F-number and shutter speed. But as it was shown later, metadata recorded by a camera is not sufficient for reliable photo quality evaluation [5, 6].

Recent works are reseaching the influence of a good photo composition [7] with identification of a foreground object and analyzing its placement inside a photo frame. As foreground objects are usually people, their position can be automatically determined using face detection algorithms.

3 PARAMETERS OF HIGH QUALITY PHOTOGRAPHS

Before we can design computer-determined features that distinguish between high and low quality photos, we must identify the criteria used by people. Extensive research was made including photography books [8, 9], scientific papers [1] as well as various photography resources available online [10]. There are numerous characteristics that define a high quality photo and since it would be impossible to implement all of them, we selected only the three most influential parameters.

The most distinguishable characteristic of a high quality photo is its *simplicity*. To satisfy the simplicity criterion, a photo should have a clear center of interest which can be identified easily (Fig. 1). Images that do not meet the simplicity criterion often include unnecessary elements that clutter the scene and make the subject recognition difficult.

Figure 1: Simplicity parameter. Due to blurred background the subject clearly stands out.

The second significant parameter that we consider is *composition*. After we identify the subject of a photo, we need to consider the interaction with other objects. The composition is a process, where we establish a sense of order for the elements within a photo. Photos taken by experienced photographers usually follow one of the established rules of composition: Golden ratio (Fig. 2), Rule of Thirds or some other rule.

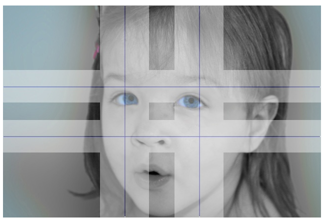

Figure 2: Composition parameter. A picture is divided into areas considering Golden ratio.

The third parameter is *color selection*. Experienced photographers often take photos in unusual lighting conditions to make the scene look appealing. Color selection of the scene is considered to be a design technique and high quality photos often contain colors that follow one of the established color schemes.

4 PROPOSED FEATURES

During our research we implemented 74 various features that analyze different aspects of a photo. Here we present only a selected few. Our approach relied heavily on identification of a photo's subject which characteristcs we compared with the rest of the photo and its background. Features that assess adherence to compositional rules such as golden ratio, rule-of-thirds and rabatment as well as the position of a subject relative to the image frame are defined. Distribution of image edges detected with an edge detector was also observed as well as photo's aspect ratio and its subject size. To assess the color palette, we calculate features that determine unique hues used in a photo along with its average hue, saturation and brightness. Simplicity of the photo was measured implicitly by comparing the color palette of the photo's subject and its background as well as with features that measure edge distribution. As a measure of visual complexity, compressed image file size is also considered.

5 DATASETS

Two different datasets were used to test our features. The first dataset was obtained by crawling recently uploaded photos on Flickr photo sharing portal. Due to the poor average quality of obtained photos, we also selected images from the Picks of the day category. The subjects of these photos were identified manually with the assistance of experienced photographers who also evaluated the composition and color balance of each photo. We obtained 258 photos, each evaluated by at least 3 different persons. For the purpose of the first experiment we used photo's overall rating which is a 3-class attribute with values: low, average and high. We excluded the photos rated as average which resulted in dataset size of 114 photos.

Obtaining a larger dataset proved to be challenging because of the time-consuming manual subject identification process. Therefore, for our second dataset we selected photos from DPChallenge portal [10] which were already evaluated. To determine also the subject automatically we decided to use only portraits so that the subjects were determined by a face detection algorithm. The photos were part of photographic contests and were rated with numerical range from 1 to 10. The average rating of a photo was 5.55 with a low standard deviation. Photos with a rating 4.5 or lower were labeled as low quality photos. Similarly, photos

Figure 3: An example of a high quality photo on the left and a low quality photo on the right. Both photos were rated by DPChallenge portal users where the photo on the left received the highest ratings in contrast to the right one which received the lowest ratings.

with a rating 6.5 or higher were labeled as high quality ones. The final dataset used in the second experiment consisted of photos previously labeled as low or high quality. It contained 1048 photos, each evaluated by at least 100 persons.

6 EXPERIMENTAL RESULTS

We evaluated the proposed features on both datasets described above. We used the Support vector machine (SVM) classifier with RBF kernel to automatically distinguish between high quality photos and low quality snapshots. We evaluate the classifier performance with a 10-fold cross validation technique and we use reliefF metric to determine feature quality [11].

In our first experiment, we calculated 73 features on 114 photos in Flickr dataset. The best classifier accuracy of 95% was achieved by using 28 best features ranked by reliefF metric. In Table 1 we show classification accuracy (in percents) which was achieved by using different number of best features (first row) ranked by reliefF metric.

Feat. used	5	10	20	28	73
CA [%]	80.2	91.3	94.7	95.3	95.3

Table 1: Flickr dataset results. Best classification accuracy of 95% was achieved using 28 automatically determined features.

In our second experiment, we calculated 71 features on 1048 photos in Flickr dataset. The best classification accuracy of 75% was achieved by using 41 best features ranked by reliefF metric. In the table below, we show classification accuracy (in per cents) which was achieved by using different number of best features (first row) ranked by reliefF metric.

Feat. used	5	10	20	30	41	71
CA [%]	69.8	71.1	72.1	72.4	74.8	73.4

Table 2: DP dataset results. Best classification accuracy of 75% was achieved using 41 features.

6.1 Feature performance

Although photos taken by experienced photographers look colorful and vibrant, the number of unique hues they contain is usually low. This was confirmed by both experiments. We observed that photos in DP dataset that contain less than 4 unique hues are preferred in comparison to more colorful ones.

Experienced photographers often crop photos to improve framing of selected subject. We observed that extreme aspect ratios are not desired. Results from DP dataset show that photos with aspect ratio from 0.81 to 1.2 are preferred.

High quality photos usually follow one of the established rules of composition. From both our datasets we observed the influence of the rule of thirds and the golden ratio. Both composition rules divide photos with two vertical and two horizontal lines. We observe that the subject center distance to nearest vertical line is the best predictor of a photo quality. Results obtained from DP dataset show that photos with a distance to nearest vertical line, defined by the rule of thirds of less than 11% of overall width are preferred (15% for golden rule). We used different templates to determine how much the photo follows the rule of thirds as well as the golden ratio rule. Application of the golden ratio rule template on a photo is shown in Figure 2.

It is not common for a subject to be placed near the photo edges. We observe that photos with the subject center distance from the left image edge of more than 38% of overall width are preferred. Similarly, distances from the bottom image edge in the interval from 23% to 86% are preferred.

We analyzed the spatial distribution of image edges detected by edge detector algorithm in order to capture the placement of distinct objects (subject). These features were ranked as the best features in DP dataset. We observe that high quality photos have the subject placed away from the image center. We also notice that the number of detected edge pixels has significant influence on perceived quality of a photo. We argue that higher number of detected edges means more distinguishable subject and/or more visually complex photo.

Complex photos require more effort and skill in order to be produced. Consequently, they are more visually pleasing to the viewer's eye. Photo complexity should be understood in terms of complexity of a photo subject and should not be mistaken for low quality snapshots that introduce a lot of background clutter. A good and simple measure of visual complexity is the size of the compressed image in jpeg format [12]. We notice that high quality photos have a higher image file size.

7 CONCLUSION

In our work, we identified significant parameters that distinguish high quality photos from low quality snapshots. On the basis of these parameters, we defined calculable features for automatic assessment of photo aesthetics using machine learning methods. Proposed features were evaluated on two different datasets. Our system was able to achieve a 95% classification accuracy on the Flickr dataset. Second dataset proved to be more challenging, photos from DP dataset were part of various photography contests and were nearing artistic photography. Automatic assessment proved to be still feasible as our classifier achieved 75% accuracy on the described dataset. Both experiments showed results that are on par or even better than in comparable work [1, 7].

Future work should focus on distinguishing between different types of photos (landscape, portrait, etc.) and their specific features. Additional research could also address the problem of automatic identification of photo subject on common photos.

REFERENCES

[1] Y. Ke, X. Tang, and F. Jing. The design of high-level features for photo quality assessment. In *IEEE Computer Vision and Pattern Recognition Conf.*, vol. 1, pp. 419 – 426, 2006.

[2] R. Datta, D. Joshi, J. Li, and J. Z. Wang. Studying aesthetics in photographic images using a computational approach. In *ECCV (3)*, pp. 288–301, 2006.

[3] R. Datta, J. Li, and J. Z. Wang. Learning the consensus on visual quality for next-generation image management. In *Proc. 15th ACM int. conf. on Multimedia*, pp. 533–536, New York, 2007.

[4] M. Boutell and J. Luo. Bayesian fusion of camera metadata cues in semantic scene classification. In CVPR 2004, pp.623–630, 2004.

[5] Y. Luo and X. Tang. Photo and video quality evaluation: Focusing on the subject. In *Proc. 10th European Conf. on Computer Vision: Part III*, p. 386–399, Springer, 2008.

[6] X. Sun, H. Yao, R. Ji, and S. Liu. Photo assessment based on computational visual attention model. In *Proc. 17th ACM int. conf. on Multimedia*, pp 541–544, New York, 2009.

[7] S. Bhattacharya, R. Sukthankar, and M. Shah. A framework for photo-quality assessment and enhancement based on visual aesthetics. In *Proc. 18th ACM int. conf. on Multimedia*, pp. 271–280, ACM, New York, 2010.

[8] M. Freeman. *Fotografov pogled : kompozicija in oblikovanje za boljše digitalne fotografije.* Tehniška založba Slovenije, Ljubljana 2011.

[9] M. Freeman. *The Digital Photography Book.* Peachpit Press, 2006.

[10] DPChallenge. http://www.dpchallenge.com

[11] M. Robnik-Šikonja, I. Kononenko. Theoretical and Empirical Analysis of ReliefF and RReliefF *Mach. Learn., pp. 23-69, Kluwer Academic Publishers, 2003.*

[12] G. Birkin. Aesthetic Complexity: Practice and Perception in Art & Design. *PhD thesis, Nottingham Trent University, 2010.*