2014 IEEE International Conference on Systems, Man, and Cybernetics
October 5-8, 2014, San Diego, CA, USA

Effective and Efficient Photo Quality Assessment

Zhe Dong
Department of Electronic Engineering and Information Science, University of Science and Technology of China
Hefei Anhui, China
Email: ustcdz@mail.ustc.edu.cn

Xinmei Tian
Department of Electronic Engineering and Information Science, University of Science and Technology of China
Hefei Anhui, China
Email: xinmei@ustc.edu.cn

***Abstract*—Automatic photo quality assessment from the perspective of visual aesthetics is a hot research topic due to its potential need in numerous applications. It tries to automatically determine whether a given image has "high" or "low" quality according to the image's visual content. Most existing researches in photo quality assessment predominantly focus on exploring hand-crafted features which may be potentially related to high-level aesthetic attributes. Most of those features are designed under the guidance of some common photography rules and prior knowledge. However, due to the subjectivity and complexity of humans' aesthetic activities, automatic image aesthetic quality assessment is very challenging. Those features are not effective enough and show varying performance on different datasets. Besides, they often require high computational cost. In this paper, we propose a set of compact aesthetic features which are not only effective but also highly efficient. We test those features on two large scale real world image datasets. The experimental results demonstrate that the proposed features achieve the best performance consistently over different datasets with a much lower computational complexity.**

I. Introduction

Photo quality assessment from the perspective of visual aesthetics has been an attractive computer vision research topic. It aims at automatically classifying images into "good" or "bad" according to their perceptual beauty. Though this task is natural to human beings, it is difficult to computers. As shown in Fig. 1, for most people it is easy to tell that images in the first row are more beautiful than images in the second row. However, it is hard to tell what specific rules that the high quality images follow. Photo quality assessment has a wide variety of applications. For example, with effective aesthetic quality assessment algorithms, image search engines can return images not only relevant but also with high quality to provide better user experience [1]. It can also help home users manage and edit their digital photos to get more attractive ones [19].

Photo quality assessment has drawn much attention in recent years. Researchers have done a lot of work and mainly focus on exploring which attributes can affect human beings' appreciation of beauty [2], [3], [4], [5], [6], [7], [8], [13]. Various aesthetic features are proposed under the guidance of subjective intuition, photography rules, and common camera settings. Though those features acquire some achievements, they usually suffer from certain disadvantages: inefficiency, redundance, or lack of robustness.

In earlier work, Tong et al. chose a "black box" model to deal with aesthetic quality assessment problem [2]. They directly concatenated many low-level features, describing image color, energy, texture, shape and so on, without careful consideration. Due to this, their method generates a very long feature vector that is 846-dimensional and leads to high computational complexity. Also its performance is not satisfactory [3] due to the serious feature redundancy problem and the lack of analysis on which low-level features are related to high-level aesthetics.

Fig. 1. For most people, it is easy to tell that images in the first row are more beautiful than images in the second row. But it is hard to tell the specific rules to the computers.

In recent works, researchers tried to analyze image aesthetics with the help of photography knowledge and intuitions [3], [4]. Datta et al. adopted a 56-dimensional feature vector to describe several high-level aesthetic attributes, such as depth of field, the rule of thirds [4]. In addition, they put more effort on exploring different image regions' relationship and extracted related features. Ke et al. conducted a deeper exploration and designed seven high-level features which had well-defined semantics such as simplicity, sharpness and exposure [3]. In [5], [6], they proposed that the subject-background relationship and image composition determined image aesthetic quality. They tried different methods to identify the photo's subject region, then extracted the contrast of clarity, brightness between subject and background as aesthetic features. Lo et al. mainly used instance-based features which required many preprocessing on the dataset to generate high/low quality template-images [13]. Different from other works, Marchesotti et al. tried a surprising solution that they applied generic image descriptors, such as bag-of-visual-words, to solve this high-level classification task [7]. Similarly, Nishiyama et al. used local color descriptors from the point of color harmony [14]. Most of these aesthetic feature extraction methods require complex process, resulting in high computational cost.

There are also some works in which experts try to cope with different types of images separately or just handle specific kind of photos. Tang et al. divided images into seven categories according to their content and adopted different assessment s-

Fig. 2. The left two images gain higher aesthetic score than the right two, since they are more colorful and people feel more harmonious.

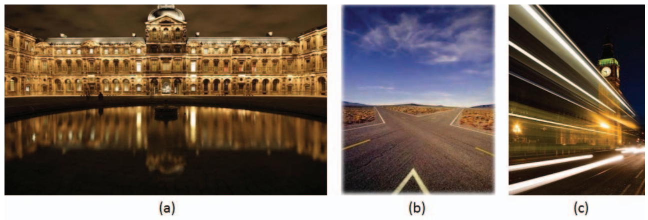

Fig. 3. Except for the rule of thirds, there still exists many other kinds of image composition rules, for example, (a) symmetry, (b) cross line, (c) diagonal. It is hard to learn all of them.

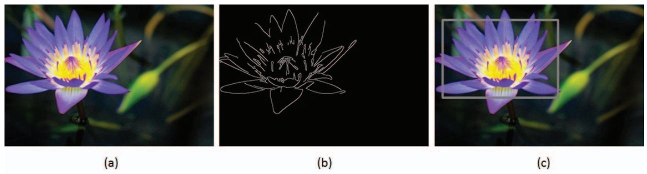

Fig. 4. An instance to show the steps of our Canny subject region detection, (a) original image, (b) Canny detection, (c) subject region marked by a rectangle.

tandards [8]. Li et al. and Male et al. restricted their processing object to images with faces [9], [10]. Similarly, Su et al. and Yin et al. dealt with scenic photos only [11], [12]. Though this can reduce the difficulty of photo quality assessment, in real applications, it is still a big problem to identify specific kind of images automatically and precisely.

In this paper, we propose a set of compact aesthetic features which are not only easy for implementation with low computational cost, but also very effective over varying datasets. Different from existing methods that are heavily constrained on the aesthetic rules, our proposed features are more flexible and as simple as possible. The experiments conducted on two popular and large scale datasets prove that our features outperform the-state-of-art methods consistently.

The rest of this paper is organized as follows. In Section II, we introduce our proposed features in details. Then the experimental results and comparisons are presented in Section III. Finally in Section IV, we give a conclusion and discuss the future work.

II. Effective and Efficient Aesthetic Feature Extraction

There exist a lot of investigations on image aesthetic quality assessment. Most of those works consider individual intuition, user interviews or photography knowledge, etc. In this section, we discuss the main criteria used by them and illustrate how we design features in details. These analyses can also provide suggestions for future research.

A. Color

Most viewers agree that photos with chromatic colors are very attractive [3], [14]. Professional photographers are serious-minded in color selection of the scene. Amateur-photographers or general users usually record what they feel interesting immediately by photos without careful arrangement. Most of their images' color schemes are lacking in creativity and impressiveness. Fig. 2 gives an example. The left two professional photos in Fig. 2 significantly impress more harmonious feelings on the viewers.

Inspired by this, researchers present many methods to describe images' color harmony. Among them, Ke et al. represented each image by a 4096-dimensional RGB color histogram [3] first. Then KNN algorithm was applied to judge whether a given test image is more likely to have high quality or low quality. The final feature was defined as the difference between numbers of good and bad images in the returned K nearest neighbors. When the training set size is very large, both the time complexity and space cost of this feature are very high. Lo et al. were inspired by this idea and tried to improve it [13]. They proposed a "color palette" method which represented the image by its five dominant colors. Then the image was reduced to a 5×3 vector which can be efficiently combined with KNN. However, the process of finding out the dominant colors for each image is not easy. It can not handle very large training set either. Similar aesthetic feature design thought can also be found in [8]. Except for inefficiency, this kind of feature has other two disadvantages. First, it just uses the K nearest neighbors' labels returned by KNN to judge whether the image is more likely to be good or bad. In this way, it loses all the original color information. Second, general KNN algorithm only calculates the Euclidean distance between two histograms as their difference. It does not consider different importance of different colors. These two weaknesses make it not that effective. Moreover, those methods heavily dependent on the training set.

Unlike extracting the color information from the global image, both Nishiyama et al. and Marchesotti et al. adopted "bag-of-color" method which used local color descriptors [14], [7]. Histograms of quantized color descriptors are taken as the images' color features. Usually this kind of histogram has thousands of dimensions and the extraction of local descriptors is inefficient. What is more, it is hard to effectively combine this high dimensional feature with other low dimensional aesthetic features.

To address above problems, we here propose a compact 12-d color histogram to describe images' color information. This 12-d color feature can preserve the color information and is also efficient for extraction. Firstly, the image is converted into the HSV color space, since HSV space is more intuitive and perceptually relevant [16]. Then we modify the non-uniform quantization method in [16] to quantize images' hue and value, as shown in (1) and (2). Non-uniform color quantization is more reasonable than uniform quantization as people have different resolving power in different ranges of light waves [16]. Here "value" information is preserved due to its relationship with image exposure condition [3], [4]. Image exposure always plays an important role in camera settings and affects the image aesthetics a lot. Finally, we traverse the quantized image pixels to form an 8-d hue histogram and a 4-d value histogram. These two histograms are directly concatenated as features $f_1 \sim f_{12}$. This method is both very efficient and effective compared with others.

$$Hue = \begin{cases} 1 & \text{if hue} \in [45, 80) \\ 2 & \text{if hue} \in [80, 140) \\ 3 & \text{if hue} \in [140, 190) \\ 4 & \text{if hue} \in [190, 255) \\ 5 & \text{if hue} \in [255, 275) \\ 6 & \text{if hue} \in [275, 320) \\ 7 & \text{if hue} \in [320, 10) \\ 8 & \text{if hue} \in [10, 45) \end{cases} \tag{1}$$

$$Value = \begin{cases} 1 & \text{if value} \in [0, 0.15) \\ 2 & \text{if value} \in [0.15, 0.4) \\ 3 & \text{if value} \in [0.4, 0.75) \\ 4 & \text{if value} \in [0.75, 1) \end{cases} \tag{2}$$

B. Subject-background separation

The foreground subject and the background are usually treated very differently in a good photo. Larger contrast between them may lead to a more impressive photo. Therefore, professional photographers will carefully set the background after determining what to shoot and skillfully control the contrast.

Luo et al. took this principle as a very important point to construct aesthetic features [5]. They adopted a complex blur detection method proposed in [17] to identify the subject region. After extracting the subject region, they calculated the clarity contrast and brightness contrast between subject and background as aesthetic features. The problem in this method is that the subject region detection step requires a lot of calculation and its performance is not satisfactory either.

Photo composition, which describes the relative location of the subject in the photo, is also proposed as an important feature for photo quality assessment in [4], [5]. It seems to make some sense according to general intuition. As reported in those papers, when observing the images, people tend to put vision emphasis on the intersections of two equally-spaced horizontal lines and two equally-spaced vertical lines. Putting the subject on one of the intersections can get better visual balance and make viewers focus interest on the photograph. This phenomenon is called "rule of thirds" in the photography field. But its importance is exaggerated. We find that many low quality images also comply with this rule. Indeed, "rule of thirds" is just a simple image composition manner which is well-known. Many high quality images use other more complex composition methods, as shown in Fig. 3. It needs lots of photography knowledge and prior experience to precisely identify what composition method the image uses. Therefore we do not model the image composition as aesthetic features, but put our emphasis on describing the subject region and the contrast between subject and background. What is more, we propose a new method to extract the subject region more efficiently.

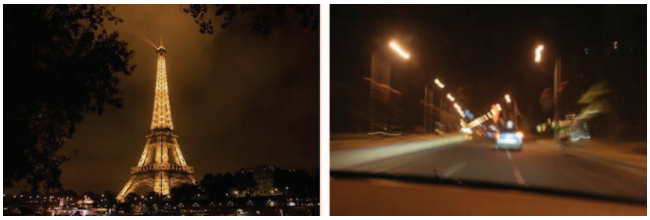

Fig. 5. In most cases, sharpness is the prerequisite for images to get a high aesthetic quality score. When taking photos, people usually discard the blur ones at once.

Fig. 6. The left image uses low depth of field to capture the bird resting on the tree with high contrast to the background. In comparison, the right image does not adjust the depth of field and therefore fails to enhance the subject.

We propose to use Canny edge detection, which is very fast, to locate the subject region. As a big difference exists between subject and background, it leads to significant edges in their boundaries. Inspired by this, we set a high threshold to filter out other less important edges. Next, a compact rectangle box which contains about 88% edge pixels is located as the subject region. An example is given in Fig. 4 to explain our subject region extraction method.

After finding out the subject region, we extract a 4-dimensional hue histogram from it as features $f_{13} \sim f_{16}$ (merge the adjacent hue section in (1), like [45,140) and [140,255)) and calculate the average value in subject region as feature f_{17}. The relative size of this subject region is used as feature f_{18},

$$f_{18} = \frac{|\mathcal{S}|}{|\mathcal{I}|} \tag{3}$$

where $|\mathcal{S}|$ and $|\mathcal{I}|$ are the subject region size and the image size.

$f_{13} \sim f_{18}$ describe characteristics of the subject. We further compute the hue and brightness differences between the subject and background as feature f_{19} and f_{20},

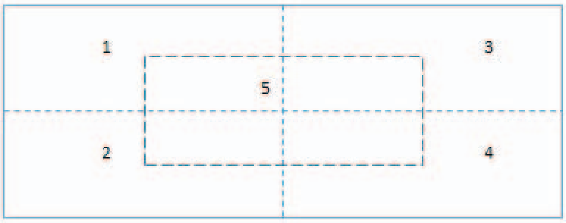

Fig. 7. Depth of filed can change image regions' sharpness. We can use this to design features to describe depth of field.

$$f_{19} = |H_{\text{subject}} - H_{\text{background}}| \quad (4)$$

$$f_{20} = |V_{\text{subject}} - V_{\text{background}}| \quad (5)$$

$H_{subject}$ and $V_{subject}$ are the average hue and brightness of pixels in subject region. $H_{background}$ and $V_{background}$ are the average hue and brightness of pixels in background.

C. Sharpness

Image sharpness is one of the most popular attributes used in aesthetic quality assessment [3], [6], [13], [14]. It usually acts as the deterministic factor. This phenomenon is easy to understand. People prefer photos with high resolution because they can easily get more details from them. High sharpness also makes the color images more gorgeous. The photo taken by a professional is extremely rare to be entirely blurry. Fig. 5 shows two images for instance. The right image which is severely blurred gets a very low aesthetic quality score. In most photo sharing websites, high resolution images can get a high score and great browsing with ease.

In view of this, features describing sharpness are essential. In [3], Ke et al. have proposed a both efficient and effective idea. They do FFT transform $\mathcal{F}$ on the gray image $\mathcal{I}$ and compute the number $|\mathcal{N}|$ of frequencies whose power is greater than a pre-defined threshold θ,

$$\mathcal{F} = FFT(\mathcal{I}) \quad (6)$$

$$\mathcal{N} = \{(u, v)||\mathcal{F}(u, v)| > \theta\} \quad (7)$$

The ratio between number $|\mathcal{N}|$ and the image size is taken as the sharpness feature. We follow this to get f_{21},

$$f_{21} = \frac{|\mathcal{N}|}{|\mathcal{I}|} \quad (8)$$

D. Depth of field

Depth of field is an important camera setting that should be considered before taking photos. Experienced photographers use low depth of field to sharpen the object they want to capture and meanwhile blur the background. This can help to emphasize the object of interest [4]. Fig. 6 shows an example that low depth of field can highlight the subject. Also the adjustment of depth of field is related with aperture settings and lenses selection that can reflect the camera's performance and user's skill. Therefore depth of field is highly related with image aesthetics.

Fig. 8. Size seems to be an unimpressive factor. But it really affects the image aesthetics a lot. In general, bigger size image with high resolution can show more details to the user.

As the depth of field can change the sharpness of different image regions, we here propose a new method to describe the depth of field. Firstly, the image is separated into five equal regions as shown in Fig. 7. Then the sharpness in each region is computed independently and therefore obtain the $\{S_1, S_2, S_3, S_4, S_5\}$. We use the maximum of region sharpness as the first depth of filed feature f_{22},

$$f_{22} = \mathbf{max}\{S_1, S_2, S_3, S_4, S_5\} \quad (9)$$

The variation of these region sharpness is taken as feature f_{23}.

E. Image Size

Image size plays an important role in aesthetic quality assessment. Though size is really unimpressive compared with other attributes, it does matter to the aesthetic perception, as demonstrated in Fig. 8. Size has the similar function to sharpness and in general, bigger size is related with higher resolution. Its effectiveness has already been proved in [4]. Chu et al. further designed a special experiment to demonstrate that the size of image can affect its aesthetic score a lot in a complex manner [15].

To derive simple and direct size-related aesthetic features, here we use f_{24}, f_{25} to describe the image size,

$$f_{24} = \mathbf{log}(\frac{X}{\omega} + 1) \quad (10)$$

$$f_{25} = \mathbf{log}(\frac{Y}{\omega} + 1) \quad (11)$$

where X and Y are the height and width of the image. ω can be set between 500 and 600 which is seen as the acceptable size threshold.

So far, we extract 25-d features for each image. All of them are carefully chosen and require no complex process. Therefore they perform very efficiently.

III. Experiments

This section verifies the effectiveness and efficiency of the proposed aesthetic features. The proposed aesthetic feature extraction method will be compared with several state-of-the-art methods [3], [4], [5], [7], [13] for photo quality assessment on two large scale benchmark datasets which are widely used.

A. Experimental setting

1) Datasets: It is not easy to collect a large scale and reliable dataset for individuals. So most experts try to download images from the photo-sharing websites to form the datasets. Here we choose two popular and large scale datasets, CUHKPQ [8] and AVA [18], for experiments.

CUHKPQ consists of images collected from professional photography websites and amateur photographers. Each image is labelled by ten independent viewers and is classified as high or low quality only if eight of ten viewers agree on its assessment. Other photos are removed. Finally, they preserve 17690 images. This filtering process makes CUHKPQ an easier dataset that "good" and "bad" images are more distinguishable. Another character of CUHKPQ is that images in it are divided into seven categories, including "animal", "architecture", "human", "landscape", "night", "plant" and "static". For each category, we evenly and randomly separate the high and low quality images into training set and test set. Finally both sets contain 8845 images.

AVA is the existing largest dataset for image aesthetic quality assessment. It contains more than 250000 images downloaded from DPChallenge.com. Many other datasets, for example the one used in [5], also consist of images downloaded from this website. However, they only contain far less numbers of images. The provider of AVA does not release the images but only their web links. We successfully downloaded 193077 of them. The links of remaining images are unable to access now due to the update of website. This website encourages users to share and score photos. Each image in this dataset has tens to hundreds scores and we take the average score to indicate the ground truth aesthetic quality of this image [3], [4], [5]. Similar to [3], [18], the top 10% and bottom 10% of the photos are assigned as high and low quality ones respectively, with the middle ambiguous images discarded. Then we randomly split them into two equal parts, one for training and the other for test.

In our opinion, AVA dataset is more close to the real Internet environment and therefore is more realistic and challenging. CUHKPQ looks more like the offline general users' collection.

2) Comparison methods: Here we implement five of the state-of-the-art aesthetic feature extraction methods as our baselines, including the 56-d features proposed by Datta et al. [4], 7-d features proposed by Ke et al. [3], 5-d features proposed by Luo et al. [5], 17-d original features proposed by Lo et al. [13] and bag-of-visual-words features of dense SIFT image descriptor with 1024 visual words proposed by Marchesotti et al. [7]. Then we take the classification accuracy as the performance indicator which is often used in related works.

SVM, which is widely used in previous photo quality assessment researches [4], [5], [6], [7], [8], is adopted for aesthetic model training. For each of the aesthetic feature extraction methods, we train a two-class SVM classifier with RBF kernel. The parameters are all automatically selected via 5-fold cross validation on the training set.

TABLE I. CLASSIFICATION ACCURACY COMPARISON BETWEEN FIVE BASELINES AND OUR PROPOSED FEATURES ON CUHKPQ DATASET.

Aesthetic Features	Accuracy (%)
Luo[5]	76.91
Lo[13]	81.76
Datta[4]	85.27
Ke[3]	81.70
Image Descriptor[7]	79.53
Proposed Features	**86.09**

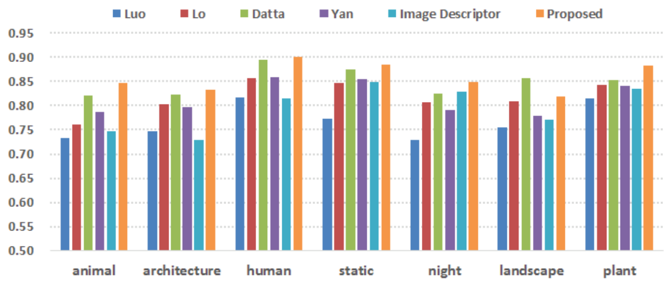

Fig. 9. The classification accuracy on the 7 categories in CUHKPQ dataset when different aesthetic features are adopted.

TABLE II. CLASSIFICATION ACCURACY COMPARISON BETWEEN FIVE BASELINES AND OUR PROPOSED FEATURES ON AVA DATASET.

Aesthetic Features	Accuracy (%)
Luo[5]	61.49
Lo[13]	68.13
Datta[4]	68.67
Ke[3]	71.06
Image Descriptor[7]	68.55
Proposed Features	**77.35**

B. Experimental results on CUHKPQ

The experimental results on CUHKPQ dataset are summarized in Table I. We can see that our proposed features outperform all other features and achieve the highest classification accuracy of 86.09%. It demonstrates the effectiveness of our proposed aesthetic features. Among the five baselines, the features proposed by Datta [4] achieve the best accuracy 85.27% which is close to ours. The features proposed by Ke [3] and Lo [13] give similar performances which are both around 81%. The Luo's features perform the worst with only 76.91% classification accuracy achieved.

Besides the overall performance, we further investigate their classification accuracy on each of the 7 categories. The results are shown in Fig. 9. It can be observed that our features are the best among the six feature extraction methods over almost all categories. The only exception is that our features perform a little worse than Datta's features on "landscape". It demonstrates that our method is not only effective but also robust.

TABLE III. RUNNING TIME COMPARISONS.

Aesthetic Features	Time (s)
Luo[5]	0.38
Lo[13]	0.96
Datta[4]	4.49
Ke[3]	0.28
Image Descriptor[7]	9.95
Proposed Features	**0.08**

C. Experimental results on AVA

Table II shows the classification accuracies of the five baselines and our proposed features for photo quality assessment on AVA dataset. Compared with Table I, the performances are much lower. As aforementioned, AVA consists of images downloaded from the Internet directly without any preprocessing. The images in this dataset are more complex and therefore this dataset is more challenging. Most of the baseline methods can't get good performances, and they show very different performances on those two datasets. For example, Datta [4] performs quite well on CUHKPQ dataset but gives a moderate performance on AVA.

Among the five baselines, Ke [3] achieves the best result of 71.06%. It is attributed to their careful feature design thoughts. Datta [4], Lo [13] and Image Descriptor [7] show similar performance (around 68.5%). Luo [5] just has a classification accuracy of 61.49%, the lowest of all. Different from those baselines, our proposed features significantly outperform them on both datasets, which demonstrates their effectiveness and robustness.

D. Efficiency comparison

Here we give the average running time of the five state-of-art methods and our proposed method in Table III. The main factor that affects the computational complexity is the image size. We randomly select hundreds of images from CUHKPQ and the average size of these images is 586×677. All the algorithms in the Table III except [4] and [7] are implemented in C++. Then all of them are tested on a Core i3, 3.4GHz, 3.4G RAM PC using Win7 operating system.

As shown in Table III, our proposed features need the shortest computing time which demonstrates their efficiency. In contrast, both the features of Datta [4] and Marchesotti [7] have much higher computational cost than other methods. This can be attributed to that features of Datta [4] need many complex image processes like Kmeans-based image segmentation. Also the general descriptor idea in [7] needs to extract lots of SIFT features from all image patches. Both of them can't perform efficiently especially when the image is large and complex. Therefore, though they are implemented in MATLAB, we can safely ensure the correctness of our conclusion.

IV. CONCLUSIONS AND FUTURE WORK

In this paper, we analyze what attributes can affect the image aesthetic quality a lot, and then propose a set of new aesthetic features which are not only effective but also efficient. Extensive experiments conducted on two large scale real world image datasets have proven the superiority of our proposed features. In the future, we will continue to explore image aesthetics at a deeper level. We will try to improve the image aesthetic quality assessment model to build a reliable automatic image rating system. And we also plan to use multi-view method such as [20] to combine more aesthetic features. We believe this will bring more amazing applications.

ACKNOWLEDGMENT

This work is supported by the NSFC under the contract No. 61201413, the Fundamental Research Funds for the Central Universities No. WK2100060007 and No. WK2100060011, the Specialized Research Fund for the Doctoral Program of Higher Education No. WJ2100060003, to Dr. Xinmei Tian.

REFERENCES

[1] B. Geng, L. Yang, C. Xu, X.-S. Hua, and S. Li. *The role of attractiveness in web image search*. In ACM MM, pages 63–72, 2011.

[2] H. Tong, M. Li, H.J. Zhang, J. H and C. Zhang. *Classification of digital photos taken by photographers or home users*. In Advances in Multimedia Information Processing, pages 198–205, 2004.

[3] Y. Ke, X. Tang, and F. Jing. *The design of high-level features for photo quality assessment*. In CVPR, pages 419–426, 2006.

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

[5] Y. Luo and X. Tang. *Photo and video quality evaluation: Focusing on the subject*. In ECCV, pages 386–399, 2008.

[6] L. K. Wong and K. L. Low. *Saliency-enhanced image aesthetics class prediction*. In ICPR, pages 997–1000, 2009.

[7] L. Marchesotti, F. Perronnin, D. Larlus, and G. Csurka. *Assessing the aesthetic quality of photographs using generic image descriptors*. In ICCV, pages 1784–1791, 2011.

[8] W. Luo, X.Wang, and X. Tang. *Content-based photo quality assessment*. In ICCV, pages 2206–2213, 2011.

[9] C. Li, A. Gallagher, A. C. Loui, and T. Chen. *Aesthetic quality assessment of consumer photos with faces*. In ICIP, pages 3221–3224, 2010.

[10] M. Males, A. Hedi and M. Grgic. *Aesthetic Quality Assessment of Headshots*. In ELMAR, 55th International Symposium, pages 89–92, 2013.

[11] H. H. Su, T. W. Chen, C. C. Kao, W. H. Hsu, and S.Y. Chen. *Scenic photo quality assessment with bag of aesthetics preserving features*. In ACM MM, pages 1213–1216, 2011.

[12] W. Yin, T. Mei, and C. W. Chen. *Assessing photo quality with geo-context and crowdsourced photos*. In Visual Communications and Image Processing, pages 1–6, 2012.

[13] K. Y. Lo, K. H. Liu, and C. S. Chen. *Assessment of photo aesthetics with efficiency*. In ICPR, pages 2186–2189, 2012.

[14] M. Nishiyama, T. Okabe, and Y. Sato. *Aesthetic quality classification of photographs based on color harmony*. In CVPR, pages 33–40, 2011.

[15] W. T. Chu, Y. K. Chen and K. T. Chen. *Size does matter: how image size affects aesthetic perception?* In ACM MM, pages 53–62, 2013.

[16] H. J. Yao. *Image Retrieval Based on Color, Space and Texture Features*. In http://www.paper.edu.cn.

[17] A. Levin. *Blind motion deblurring using image statistics*. In NIPS, pages 2–4, 2006.

[18] N. Murray, L. Marchesotti, and F. Perronnin. *AVA: A largescale database for aesthetic visual analysis*. In CVPR, pages 2408–2415, 2012.

[19] S. Bhattacharya, R. Sukthankar, and M. Shah. *A Framework for Photo-Quality Assessment and Enhancement based on Visual Aesthetics*. In MM, pages 271–280, 2010.

[20] J. Yu, Y. Rui, Y. Tang, D. Tao. *High-Order Distance-Based Multiview Stochastic Learning in Image Classification*. In IEEE Trans. Cybernet, 2014.