

Aesthetic Quality Assessment of Photographic Images: A Literature Survey

Sruthy Suran
Department of CSE
College of Engineering Poonjar
Kottayam Kerala, India

Sreekumar K.
Department of CSE
College of Engineering Poonjar
Kottayam Kerala, India

ABSTRACT

Aesthetics is a branch of philosophy which deals with the study of emotions in relation to the sense of beauty. Nowadays, there is a tremendous increase in the use of digital images as a means for representing and communicating information. With the considerable increase of consumer photos, evaluating the quality of photos has become a difficult task. People are more interested in photos that are visually pleasing. The aesthetic beauty of a picture is determined using aesthetic quality factors like prettiness, cuteness, neatness, cuddliness, loveliness etc. Aesthetic quality assessment is a challenging task that require understanding of subjective notions . Aesthetic quality score of an image can be calculated using low level features such as contrast, sharpness, colorfulness etc. This paper provide a survey of aesthetic quality assessment of photographic images and a brief description of existing approaches.

General Terms

Image Aesthetics, Image Quality, Photo Quality Assessment

Keywords

Aesthetic attributes, Aesthetic score , Feature extraction, Image aesthetic assessment

1. INTRODUCTION

With the increase in number of digital cameras, there is an explosion in the number of digital images created. Photos are an important means to preserve the memory of important events in one's life. They capture a baby's first developments, a picnic, wedding etc. To remember and share these events, people create visual compositions such as collages, calendars, or photo books. Image quality assessment aims to classify the photographs into high quality or low quality automatically. Photography can be defined as the art or practice of taking and processing photographs. Aesthetics in photography means how people characterize beauty in this form of art. The photographic images that are soothing to the eyes are considered to be higher in terms of their aesthetic beauty. A normal person may be interested in how soothing a picture is to the eyes, while an artist may be looking at the composition of the picture, use of colors, light and shade etc, and a professional photographer may be looking at the sharpness or color contrast of the picture, or whether the "rules of thumb" in photography have been maintained. All these issues make the measurement of aesthetics in photographs extremely subjective.

Aesthetics can be formally interpreted as an apparently simple matter to what is beautiful, few can meaningfully communicate the definition of aesthetics or how to attain a high level of aesthetic quality in photographs. There is a need to have a formal or mathematical clarification of aesthetics in photographs. It is broadly believed and can often be

experimentally verified that aesthetics is at times very subjective. That is, the same photograph can be valued by some viewers but not by certain others. The "taste" and sophistication of the viewer often determines the aesthetic rating specified by the viewer. For years, Photo.net is used by photographers to rate the photos. Here a photo is rated with two dimensions, ie aesthetics and originality, each with a score between one and seven. Photo.net explains that the score of one means ugly and the score of seven means beautiful. According to statistics by Flicker, an average of 6.5 million photographs are uploaded daily by its users. Thus, there is a great need for multimedia applications to manage, rate and edit such content.

2. IMAGE QUALITY FEATURES

To determine photo's quality, color and composition are the key factors, and they are highly related to human's perception

2.1 Color Palette(CP)

It is a finite set of colors used for the management of digital images. It is important to estimate a camera's color response such as its color shifts, saturation, and the effectiveness of its white balance algorithms. One important image quality factor is color accuracy. Most of the viewers choose enhanced color saturation and the most accurate color need not be the most pleasing.

2.2 Layout Composition(LC)

Arrangement of visual elements in an image is defined as layout composition. Instead of using traditional rule specific methods like rule of thirds and visual balance, template based principle is used. They are obtained for the H,S,V channels. Let $L1$ be the distance between the input image and dH and dL be the templates for high quality and low quality images. LC features $f2$ to $f5$ is obtained from the vale $dL - dH$.

2.3 Texture

Refers to characteristics of surface and appearance of an object given by the size, shape, density, arrangement, proportion of its elementary parts. Texture measures look for visual patterns in images and provide information about the spatial arrangement of color or intensities of an image. Image texture can be found in natural scenes captured in an image or can be artificially created. It is useful in segmentation and classification of images. Structured approach and Statistical approach is used for analyzing image texture. In structured approach a set of primitive texels in regular or repeated pattern is considered and in statistical approach a quantitative measure of the arrangement of intensities in a region is considered.

2.4 Contrast

It is the difference between darkest and brightest areas of an image that makes an object unique. With increase in contrast we can increase the separation between darkest and brightest part, ie making shadows darker and highlights brighter. By decreasing contrast we bring the shadows up and highlights down to make them closer to one another. Increasing brightness usually makes an image more vibrant and decreasing brightness usually makes an image dull. Also when contrast is very high it leads to loss of details in an image.

2.5 Sharpness

Can be defined as the contrast along edges in a photo. It is defined as the boundaries between different colors or tones and it also gives the amount of details an image can convey. By using rise distance of the edge, sharpness of an image can be measured. Rise distance is a good indicator of image sharpness.

2.6 Content

Means any information that can be derived from an image, ie what emotions an image express.

2.7 Composition

It is the arrangement of relative object and elements which suits the core idea of the work. It is the key feature of a good work of art. Focal length, aperture, angle of position of the camera also greatly affects composition. A good composition is a method for guiding the viewers eye towards the most important features of an image. Layout composition and edge composition are used in images.

2.8 Distortion

It is defined as a deviation from rectilinear projection. It can be irregular or can follow many patterns. It affects the quality of images and it is commonly visible in low cost cameras, mobile phone cameras etc. Most commonly found distortion is radially symmetric distortion. Radial distortion can be classified as Barrel distortion, Pincushion distortion and Mustache distortion. In barrel distortion the image magnification decreases with distance from the optical axis and in pincushion distortion the image magnification increases with distance from the optimal axis. Mustache distortion can be considered as a mixture of both.

2.9 Noise

It is visible as grain in film and pixel level variations in digital images which is caused due to variation in digital images. It arises due to the effects of the photon nature of light and thermal energy of heat inside image sensors and amplifiers. Visual noise is used to estimate the visual strength of noise using the noise spectrum. Noise scales powerfully with the size of the pixel. It is low in digital SLR's with a pixel of atleast 4 microns square and it gets ugly in compact digital cameras and camera phones containing small sensors. Software noise reduction can be used to reduce noise by smoothing the image.

2.10 Color moiré

In images repetitive patterns of high spatial frequencies like fabrics and picket fences appear due to artificial color banding. It mainly occur due to lens sharpness, demosaicing software and anti-aliasing. Color moire is measured by Log Frequency, where a sine chart is used consisting of logarithmically increasing spatial frequency.

2.11 Lens flare

It occurs due to the reflection of light between the lens elements and the camera barrel. It also create fogging effect in images. It is formed due to internal reflection and scattering from material inhomogenities in the lens. Flare can be defined in two ways: as visible artifacts, and as a haze across the image. A lens flare is sometimes added to an artificial image in order to add a sense of realism to show that the image is an unedited original photograph.

2.12 Lateral Chromatic aberration (LCA)

In some cases due to the aberration of lenses the color of the image gets a purple tinge. It is also called color fringes that causes colors to focus at different distances from the image centre. It is mostly visible near image corners. LCA can be measured using SFR, SFRplus, eSFR ISO, SFRreg and Checkerboard using tangential edges near the image boundaries.

2.13 Color Accuracy

It is an important image quality factor which is relevant in medical and technical photography. Color accuracy is measured against standard chart reference value (CSV) containing measured color values which can be changed to satisfy consumer preferences. Color accuracy is affected by RAW converter, signal processing and by Bayer color filter array.

3. AESTHETIC QUALITY ASSESSMENT AND SCORE PREDICTION

The aim of image quality assessment is to provide computational models that can automatically predict image quality. Pattern recognition tools, db and textual image quality tools are some integrated collection of tools used for aesthetic quality assessment. Direct and indirect methods are used to assess the quality of an image. It include subjective and objective methods. In direct method, subjective method include psychophysical experiments with the participation of human experts and objective method include computing suitable image qualities directly from the digital image. In indirect method, objective method include quantifying the performance of an image-based task done by the domain expert and/or by a computational system, and subjective method include the performance of the imaging /rendering device on suitable set of target image is assessed by using ad-hoc designed software tools one or more direct method.

For an image with some imperfections quality can be defined by the absence of certain type of distortions like ringing, compression, banding and blur. No reference (NR) type of quality assessment is significantly more challenging because such quality metrics do not utilize a reference image, they only look at the image under test. Full-reference (FR) type of image quality assessment perform a direct comparison between the image under test and the reference image.

Golden triangle rule is the composition rule in photography where a diagonal line divide the image corner to corner. This feature can be calculated by creating three diagonal guides by marking the vertical edges of the screen 1/6th of the height from the top and 1/6th from the bottom. Similarly mark the horizontal edges 1/6th of the width from the left and 1/6th from the right. Now connect the upper left horizontal mark with the lower right vertical mark, the upper left vertical mark with the lower right horizontal mark, and the upper left corner

with the lower right corner. The picture can be made more pleasing and dynamic by placing natural elements that form a line along these diagonal guides.

Exposure of lights can be considered as a good discriminant between high and low quality photos. Too much exposure leads to brighter shots which results in lower quality photos and those that are too dark are often also not interesting.

Rule of thirds is the composition rule in photography which dictate how an image should be composed in order to create an aesthetically pleasing result.. This rule states that the centre of interest in a photograph is situated in one of the four possible intersections of the lines that divide the image into nine rectangle. Photographs that follow this rule have the main object stretch from an intersection up to the center of the image. The rule of thirds is applied by aligning a subject with the guide lines and their intersection points, placing the horizon on the top or bottom line, or allowing linear features in the image to flow from section to section. The main reason for observing the rule of thirds is to discourage placement of the subject at the center, or prevent a horizon from appearing to divide the picture in half.

4. RELATED WORKS

Kuo et al[8] proposed a method which is implemented and visualized in online aesthetic assessment systems. This method is designed for a set of features that follows an instance based approach. Here classification accuracy(ACC) is used for the evaluation of classification performance systems. Image features are based on color composition and spatial composition such as color palette, layout/edge composition, and global texture features etc. Unbalanced dataset is measured using the area under the ROC curve. Experimental results show that, to achieve better classification on different variety of photos , use of simple technique is recommended. The main reason for observing the rule of thirds is to discourage placement of the subject at the center, or prevent a horizon from appearing to divide the picture in half.

Xiaoou et al[2] proposed a method based on photo quality assessment frame work. Here both regional and global features are selected and combined to assess photos in different categories. Inorder to extract subject areas three methods are used, which include clarity based region detection, layout based region detection and human based detection. Features focused on are based on dark channel feature, hue composition feature and human based features. SVM classifier is used for combining all the features. Experimental results show that the effectiveness of the method highly depends on the visual content.

Christel et al [3] proposed a method for designing a harmony quality assessment method to define what is visible or not in images and videos. It is based on a no reference perceptual harmony-guided quality map as well as a score of disharmony. Visual masking is applied to estimate the extent to which an image area can be perceived disharmonious. Color rules are used for the detection of disharmonious image parts. Visual masking is applied to estimate the extend to which an image area can be perceived disharmonious.

Tsung et al[4] proposed a model based on multi method fusion. Machine learning tool used for fusion of multiple methods is support vector regression approach. The scores obtained from multiple quality indices is integrated into one score. A fused IQA method selection algorithm called Biggest Index Ranking Difference (BIRD) is developed to reduce the complexity of multi method fusion(MMF). SFMS and BIRD

algorithms are used to reduce the number of fused methods. Inorder to achieve satisfactory performance three methods are needed for the fusion and the computational time is around one minute per image. Experimental result show that the approach performs better than well-known image quality indices.

Anish et al[5] focus on an IQA model called Natural Image Quality Evaluator(NIQE). Here a Multivariate Gaussian (MVG) model is used, in which a collection of quality of aware features are integrated. Quality aware features are based on a simple space domain called Natural Scene Static(NSS)model. The quality of the distorted image can be given as a distance metric between the model statistics and those of the distorted image. It was shown by experimental results that this model outperforms FR IQA models and competes with top performing NR IQA trained on human judgment of known distorted images.

Luca et al [6] proposed a method based on generic image descriptor. For aesthetic assessment a generic content –based local image signature is used. Descriptors used are computed from low level local features. Bag-of-Visual-Words and fisher vector are two families of generic image descriptors used for encoding the distribution of aesthetic properties. Gradient information is encoded using SIFT and color information. GIST descriptor is also used for scene categorization.

Hsiao et al [7] proposed a method based on bottom up aesthetic modeling. This method is used for scenic photograph and uses an aesthetic library called bag-of –aesthetics preserving (BoAP) feature image representation. Inorder to overcome both implicit and explicit aesthetic features, the BoAP features are extracted from different feature spaces. Experimental results shows that this method gives an accuracy of 92.06% Among the selected features in the model contrast features are more effective than absolute features .This method performs better than state–of–the–art rule specific methods. It also overcome both implicit and explicit features with learning process.

Xin et al [1] focus on a method based on deep neural network approach. For a given input image it compute an automatic learning feature of the image. RAPID system is used in this approach. Double column deep convolutional neural network approach is used for doing aesthetic quality rating and categorization. Among the selected features in the model contrast features are more effective than absolute features. This method performs better than state–of–the–art rule specific methods approach is used for aesthetic quality rating and categorization. Here the main idea is to combine heterogeneous inputs generated from the image which include a global view and a local view, and to unify them.

Luming Zhang, et al [9] introduce a probabilistic mode for photo aesthetics evaluation by discovering both global and local spatial structure among image region. In this method graphlets are first extracted which represent photo local composition. Graphlets are then projected into the Grassmann manifold, based on which a manifold embedding algorithm encodes global layout and multi-channel visual features into graphlets. These post-embedding graphlets are finally integrated to form a probabilistic measure for evaluating photo aesthetics.

Cong cong et al[10] focus on a machine learning scheme to determine aesthetic visual quality of digital images of paintings. Feature extraction is based on color and composition. Composition features are determined through

the analysis on shapes and spatial relationship of different parts of the image and color features are based on HSL space. About 40 features are proposed and a feature set is formed. In this approach global and local features are considered.. Global features are measured by classic models and local features are measured by a specific metrics.

Tunc, Ozan Aydin et al[11] present a system for automatic aesthetic evaluation of photographic images. This approach is based on the concept of the no-reference image quality assessment. The main aim is to focus on rating image aesthetic attribute rather than detecting image distortions. This system compute calibrated ratings for a set of fundamental and meaningful aesthetic attributes which forms an aesthetic signature. It introduce an experimental procedure for obtaining subjective ratings for each aesthetic attribute, and offered a set of metrics capable of perfectly predicting these subjective ratings. This method outperforms the current state-

of-the-art in predicting an overall aesthetics rating, as well as introducing novel application areas.

5. EVALUATING PERFORMANCE

For evaluation a set of images are considered, these images are provided with a score by human users. The score scale is a set 0 to 10, where higher score indicate better quality and low score indicate low quality. Multiple users independently rate on the same image. The aesthetic score of an image calculated as the mean value of all submitted scores. The score given by human users are then compared with the score given by computers. Dataset for evaluation are taken from websites like DPchallenge, Photo.net, Terragalleria, Alipr etc. These dataset contain images and aesthetic judgments received from members of the community. Collecting ground truth data in this manner is inexpensive and expedient way to obtain aesthetic judgments from multiple individuals.

Table 1. Comparison of Image Aesthetics Assessment Methods

Method	Author	Year of Publication	Accuracy	Advantages/Disadvantages	Learning Method
RAPID	Xin Lu,Zhe Lin, Jian chaoYang, James Z	2014	71.20%	i)It take more computation time. ii)local and g lobal features of an image is taken into account.	Double column Deep Convolutional Neural Network
Data – driven machine learning Problem	Congcong Li,Tsuhuan Chen	2013	87%	i) regression approach is used to estimate the quality score	Naïve Baye and Adaptive Boosting Classifier
NIQE(Natural Image Quality Evaluator)	Anish Mittal,Rajiv Soundararajan and Alan C Bovik	2012	91.47%	i)NIQE performs better than the FR PSNR and SSIM	Not applicable
Generic image descriptor	Luca Marchesotti, Florent Perronnin, DianeLarlus, Gabriela Csurka	2011	ACC –86% AUC-93%	i) Acceptable computation time for real-time applications. ii) Classification rates are better. iii)Blur information can also be added to the color descriptors	Linear SVM
Bottom –up aesthetic modeling method	Hsiao Hang,Su, Tse-wei , chieh-chikao,WinstonHsu ,shao-YiChien	2011	92.06%	i)multiple scores can be given ii)computation process is highly efficient iii)contrast information is considered	Bag-of – Aesthetics Preserving
MMF	Tsung-Jung Weisi Lin, C.C.Jay Kuo	2013	91.07%	i)highly flexible ii) excellent performance	SVR

6. CONCLUSION

This paper gives a survey of aesthetic quality classification of photographs which provide an overview of need for image aesthetic quality. A set of visual features related to characteristics of image quality and aesthetic values are used for aesthetic estimation. The image quality assessment and existing approaches are also discussed in this paper. There are different methods for feature extraction and classification. Neural network approach, double column deep convolutional neural network, naïve bayes and adaptive boosting classifier are some of the classification techniques used. As a future work multiple aesthetic quality score (excellent, good, average, bad) can be given instead of finding quality score as binary (good or bad).

7. REFERENCES

- [1] Xin Lu, Zhe Lin, Jian chao Yang, James Z wang, "RAPID: RAting Pictorial aesthetic using Deep learning", ACM 2014
- [2] Xiaou Tang, Wei Luo and Xiaogang Wang, "content-Based photo quality assessment", IEEE 2013.
- [3] Christel Chamaret and Fabrice Urban, "NoReference Harmony-guided Quality Assessment", IEEE Conference on Computer Vision And Pattern Recognition Workshops 2013.
- [4] Tsung-Jung Liu, Weisi Lin, C.-C. Jay Kuo, "image quality assessment using multi method fusion", IEEE Transactions on Image processing, VOL.22, No.5, May 2013
- [5] Anish Mittal, Rajiv Soundararajan and Alan C Bovik, "Making a completely blind image quality analyzer", IEEE 2012.
- [6] Luca Marchesotti, Florent Perronnin, Diane Larlus, Gabriela Csurka, "Assessing the aesthetic quality of photographs using generic image descriptors", IEEE international conference on computer vision 2011.
- [7] Hsiao-Hang Su, Tse-wei chen, chieh-chi kao, Winston H Hsu, shao-Yi Chien, "Scenic photo quality assessment with bag of aesthetic-preserving features", ACM 2011
- [8] Kuo-Yen Lo, Keng-Hao Liu and Chu-Song Chen, "assessment of Photo Aesthetic with Efficiency", ACM 2010.
- [9] Luming Zhang, Yue Gao, Roger Zimmermann, Qi Tian, "Fusion of Multichannel Local and Global Structural Cues for Photo Aesthetics Evaluation", IEEE 2014.
- [10] Congcong Li, Tsuhan Chen, Aesthetic Visual Quality Assessment of Paintings IJSTSP Tunc, Ozan Aydin, Aljoscha Smolic, and Markus Gross, "Automated Aesthetic Analysis of Photographic Images", IEEE 2015.