

Image Aesthetic Quality Assessment and Enhancement

Faiji Ansari¹ Akshaya Sharma² Mohit Jawalgekar³ Kuldeep Awachat⁴ Prof. M.V.Phatak⁵

^{1,2,3,4}Student ⁵Assistant Professor

^{1,2,3,4,5}Department of Computer Engineering

^{1,2,3,4,5}M.I.T Pune

Abstract— Aesthetics refers to the principle of the nature and appreciation of beauty. Judging beauty and other aesthetic qualities of photographs is highly subjective task. This system classifies the pleasing and non-pleasing images based on photography rules ex.(rule-of-third, depth-of-field, rule-of-simplicity). The proposed system automatically infers aesthetic quality of images by finding the aesthetic score using regression technique. Based on threshold value the image is classified as pleasing and not pleasing image. Then system will check which photography rule are not obeyed by the image ex.(rule-of third, depth-of-field, rule-of-simplicity) and try to enhance image based on these rule so the image will at-least cross the threshold value. For all non-pleasing image this system will automatically enhance and improve the aesthetic quality to make it a pleasing image. Application of this system includes digital photography, still advertisement, educational institutes and for designing attractive website. This system can become a guide for people if it is integrated with a smart camera. It also helps in improving the image quality in applications like image-viewer, image-editor and also you-tube images. It can also be used in flash cards which helps student to learn faster and in an interesting manner.

Key words: Aesthetics, Enhancement, Score Prediction, Quality Assessment, Regression

I. INTRODUCTION

A photograph which has visually dominant subjects it normally induces stronger aesthetic interest. In the study of photography and aesthetics, photographers deliberately avoid uniform sharpness of focus and illumination as an approach to achieve higher image aesthetics[1]. This approach is based on the basis that our eyes are attracted to salient elements that are acutely sharp, bright or colourful in images. However, due to the limited selective focusing capability of the unfavourable lighting, and the infeasibility to exclude disturbing background, makes it difficult, especially for casual photographers, for making the main subjects visually salient. This results in prolonged searching for the subjects in a photograph which leads to reduce

satisfaction of viewing and thus, poorer aesthetics experience. In this paper, the aesthetics of the input image is enhanced with the approach to enhance image aesthetics using regression methodology. The key idea is to alter image features of the objects in the photograph using the visual content as a machine learning problem. The goal is to produce a improved-aesthetic version of the input image that can grab the viewer's focus to the most important objects in the image, which makes these objects the main subjects. This is useful for enhancing photographs that do not have any apparent main subjects, or for photographs that one wishes to swap the role of the main subject with some other objects.

The key idea is to alter image features of the objects in the photograph using the visual content as a machine learning problem. The goal is to produce a improved-aesthetic version of the input image that can grab the viewer's focus to the most important objects in the image, which makes these objects the main subjects. This is useful for enhancing photographs that do not have any apparent main subjects, or for photographs that one wishes to swap the role of the main subject with some other objects.

II. PREVIOUS WORK

Our work is related to the three main areas of the research i.e. estimating visual attributes, estimating the aesthetics of photographs, and enhancement of aesthetics.

One of the popular systems which was present in the market is:

ACQUINE - Aesthetic Quality Inference Engine, a publicly accessible system where users upload their photographs and have them rated automatically for aesthetic quality. The system uses a support vector machine based classifier which extracts visual features on the fly and performs real-time classification and prediction.

The disadvantage of ACQUINE is that the developers of the system have not disclosed the algorithms or features considered while classification.

But now its link is non-operatable.

Paper	Author	Description
Algorithmic Inferencing of Aesthetics and Emotion in the Natural Images : An Exposition[1]	Ritendra Datta, Jia Li, and James Z. Wang	Finding out the basic rules of photography like, 1.Rule of Third 2.Depth of Field 3.Rule of Simplicity 4.Diagonal Dominance 5.Visual Balance
Optimizing photo Composition	Ligang Liu Renjie Chen	Predicting the aesthetic score of an image using regression techniques.
Rule of Thirds Detection from the Photograph	Long Mai, Hoang Le, Yuzhen Niu, and Feng Liu	Identify the aesthetic quality of images according to rule of third technique of photography.
Beauty is Here: Evaluating Aesthetics in Videos Using Multimodal Features and the Free Training Data.	Yanran Wang, Qi Dai, Rui Feng, Yu-Gang Jiang.	Finding the aesthetic score through composition rules.

Table 1: Papers & its Description

III. PROPOSED WORK

We propose a system which will work on the aesthetic of the image to make it aesthetically pleasing image. For this the system does object detection using saliency analysis. Then the system uses certain popular rules of photography which are used as a base for classification by the classifier of image and it automatically infer the aesthetic quality of images. It then predicts the score of the image. And then enhances their aesthetic value using their visual content.

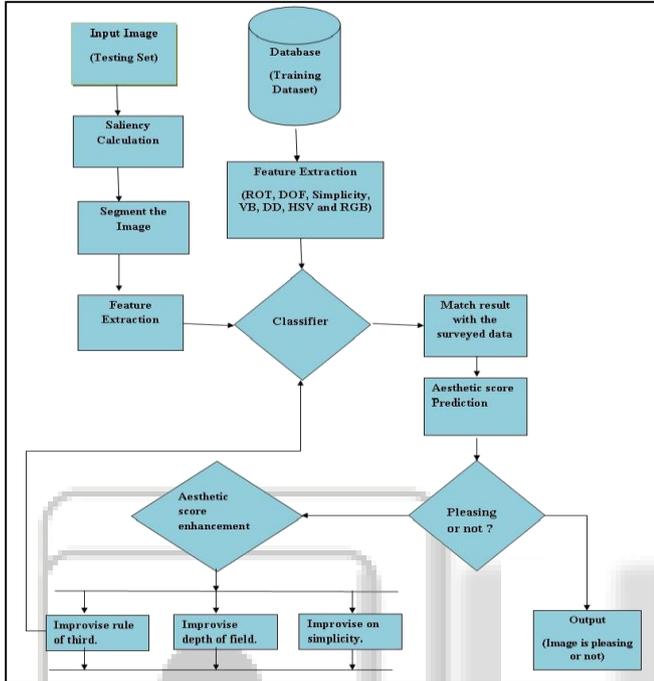


Fig. 1: Figure Represent The System Architecture Diagram of The System.

A. Saliency Analysis

Inspired by the success of using low-level saliency as alternative in other multimedia applications, such as multimedia retargeting, our method explores the saliency analysis to detect important content. Visual saliency measures the low-level stimuli to the enhance human visual system. A variety of methods have been recently developed to estimate the visual saliency from images. Our method selects the Achanta’s approach to saliency detection [8].

In this method for finding salient regions a contrast determination filter is used that operates at various scales to generate saliency maps containing “saliency values” per pixel.

Combining this individual maps result in the final saliency map. The original image and its saliency map is shown in the fig. 2

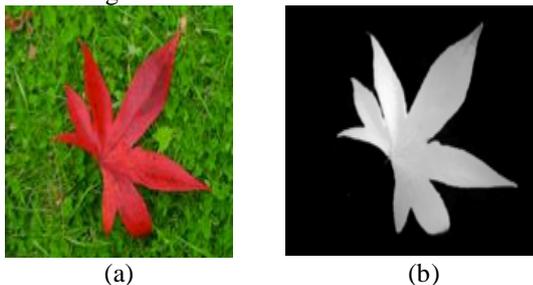


Fig. 2(a): Original Image Fig. 2(b): Saliency map

B. Classifier

The classifier classifies the image based on the rules of photography i.e whether the image is following that particular rule or not. Rules mainly focus on three rules as follows:

1) Rule of Thirds:

The Rule of Thirds is one of the most important and effective composition rule used by professional photographers for the creation of high-quality photos. This rule indicates that an image should be imagined as divided into nine equal parts by dividing two equally spaced horizontal lines and vertical lines. The most important compositional elements should be placed along these lines or their intersections.[3,6]

However instead of using the centroid of the whole saliency map for finding the objects location, in this the method finds a minimal rectangle that contains at least λ of the total saliency value as follows. Algorithm is like:

$$p_c = \frac{\sum_{p \in W} (w_p \times p)}{\sum_{p \in W} w_p}$$

where

p_c = Two element vector containing x and y coordinate(pixel with max. saliency value).

W =Block with maximum saliency value.

w_p =Saliency value at pixel p in block W.

p = location of pixel p.



Fig. 3(a): Follows ROT Fig. 3(b): Does not follow ROT

2) Depth of Field:

In particularly as it relates to film and photography, depth of field (DOF), also called focus range is the distance between the foreground and background objects in a scene that appear acceptably sharp in an image. A lens can particularly focus at only one distance at a time, the decrease in sharpness is gradual on all side of the focused distances, so that within the DOF, the unsharpness is imperceptible under normal viewing conditions.[7]

The process for DOF computation is performed by features extraction with the help of overall blur-ness, contrast .Overall blurriness contrast is the saliency map which is used for the colour based segmentation of the image. The contrast is calculated as the difference between the blur factors of the foreground and the background.



Fig. 4(a): Follows DOF Fig. 4(b): Non-DOF

3) Simplicity

Place the subject against a neutral background as the background can draw viewers’ attention to the subject of

interest in a photograph. Backgrounds can be entirely neutral, like a solid backdrop or a cloudless sky.

The more technical method of achieving simplicity involves focusing on the subject while ensuring the background is unfocused.

The picture taken to represent simplicity should have a clear reason that object was chosen.

This related to Shallow Depth of field which blurs out the background and focuses on what is on the middle.

The procedure for simplicity is, first find the number of the background blocks using the number of background blocks = 225-(No of blocks which satisfy the compactness criteria).Then compute the background simplicity using (Considering 8-neighbourhood)[4,9]

$$d(b_i, b_j) = e^{-\frac{||X(b_i) - X(b_j)||^2}{\sigma_a}} e^{-\frac{||F(b_i) - F(b_j)||^2}{\sigma_c}}$$

$$\text{where } F(b) = [v, v.s. \sin(h), v.s. \cos(h)]$$

b=block in background B.

h,s,v = hue, saturation and value for the average colour of the block b in background B.

$$F_{\text{simplicity}} = (1/|B|) \sum d(b_i, b_j)$$

where $b_i, b_j \in B$



Fig. 5(a)



Fig. 5(b)

Fig. 5(a): Simplicity following Fig. 5(b): Simplicity not following.

C. Score Prediction

The aesthetic score function is defined as a combination of the above aesthetic measurement scores:-

$$E_a = \frac{\omega_{RT} E_{RT} + \omega_{DA} E_{DA} + \omega_{VB} E_{VB}}{\omega_{RT} + \omega_{DA} + \omega_{VB}}$$

Where $\omega_{RT} = 1$ and $\omega_{VB} = 0.3$ are fixed weights. ω_{DA} is 1 if there are detected diagonal lines in the image, otherwise it sets to 0. Salient-regions sizes, According to the three aesthetic guidelines is superior to using just one rule e.g., the rule of thirds, it means that this combined score is not restrictive enough. Considering a simple example which contains only single salient object, this object can be placed on the power-points of the image (according to rule of thirds). That is, there are so many cropping frames that have equal highest scores. We now introduce the region size score which plays an important rule in stabilizing the optimization problem by eliminating much of this freedom.



Fig. 6(a): 3.0



Fig. 6(b): 4.5

Fig. 6(a): original image. Fig. 6(b): enhanced image

The example of the score prediction is shown in the fig.6. The number present below the images are the score of that particular image. The fig.6(a) is the original

image. This image does not follow the rule of third, So the score fetched by it is less making it a non-pleasing image. The image in the fig.6.(b) is the modified image. In this the rule of third is getting followed by the image which makes it more pleasing and Hence increases the score of the image.

D. Enhancement:

The enhancement module's functionality is to improve the aesthetic value of the give image. This enhancement engine is different from the other image quality improvising applications in the way that here the enhancement is done base on the photography rules. These rules are universally true and applicable. [2,10] The image is turned in to the pleasing image according to the photography rules. The enhancement is done on the non-pleasing images only. The score and the result of the classifier decides whether to enhance the image quality or not. The enhancement is done for the rule which is not following. The classifier gives the result of following and non-following amongst the rule of third, depth of field and simplicity. For the non-following rule changes are made accordingly. To improvise the rule of third the image is cropped and to improve depth of field [8] the background is made blur and the recognised object is kept in the focus. To improvise over the simplicity the background is turned even and plan. the focus.

1) Rule of Third:

The rule of third is based on placing the detected object on or near the four power points identified for the image. The image's object if placed on these point then it follows rule of third. Or else if the image's object is placed else were like the centre of the image or away from the four power point if do not follow the rule of third. When a particular image is identified as not following the rule of third then it is made to follow it by cropping it. The image is cropped in such a way, So that the object of the image comes over two of the power point of the over-all image. For this purpose the image is cropped using the hard cropping technique. There are two method of cropping the image, that is soft cropping in which purely resizing the image. In this the image is required to fit the desired dimension of the box, and if the dimension specified is not making the image fit properly. Then image is fit by laying of some blank space in the box specified making the entire image visible without cutting any of its part. In this paper of rule of third enhancement the technique used for cropping is the hard cropping technique. In hard cropping the image is filled in the specified dimension and the part of image which is not fit in desired dimension is cut out.

The example of the rule of third enhancer is in the fig.6. This figure shows the image fig.6.(a) is the original image which does not follow the rule of third, So it is having less aesthetic score. Image in fig.6.(b) is the modified image by performing the hard cropping which make the image to follow the rule of third and Hence increases the score.

2) Depth of Field:

For enhancing the image in contest of depth of field. The noise is induced in the background and reduce details of the image. With this the object seems to be sharpened. The noise added in the background of the image gets the object in the focus and more eye catching to the viewer. Gaussian Blur transformation acts on each pixel of the image, then it sets its value to the average of all the pixel values present in

a radius defined in the dialog. The higher the value, higher is the amount of blur. The formula to calculate transformation in one dimension is:-

$$G(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}}$$

The formula for calculating transformation value in two dimension is;

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \quad [11][12]$$

Where x = distance from the origin in the horizontal axis
 y = distance from the origin in the vertical axis and
 σ = standard deviation of the Gaussian distribution.

There are various filters present for noise insertion. The Gaussian filter is one of the filters which is used to blur the image. The Gaussian function is used to blur the background of the image by detecting the object of the image.

IV. CONCLUSION

In this project we have seen some key aspects of image processing have been applied to classify image having single object. It extracts the features of images by applying five most prominent rules of photography. The classifier is trained which categorizes image as pleasing or not. If the quality of image is not above the threshold value then the image is considered as not pleasing and it modifies the aesthetic parameters which make images pleasing. The proposed system can be used in still advertisement to promote the product by making the advertisement attractive. It can also be used in attractive website designing and used in educational institutes taking into consideration the visual aspects of images. It can also be used for designing flash cards to make them more attractive for children by checking its aesthetics. This system can become a guide for people if it is integrated with a smart camera. Our work is a significant step towards the highly challenging task of understanding the correlation of human perception of the pictures they see by a computational approach. We hope that this system will be helpful to foster image classification.

V. FUTURE WORK

The system can be enhanced for applications in the image re-targeting and video compression.

The proposed system is a prototype of automatic aesthetic classification systems and can be further extended to build large systems.

The system can be extended to include gender preferences based on objects in image.

The system can be extended for multiple object images.

Implement IAQE@E system on videos.

Implementation of new rule coming in the photography world.

ACKNOWLEDGMENT

We would like to acknowledge and would express our special thanks of gratitude to our honourable project guide Prof.M.V.Phatak, of Department of Computer Engineering, Maharashtra Institute of Technology, Pune for guidance and encouragement in carrying out our work. The supervision and support that she gave truly helped the progression and

smoothness of the project. The co-operation is much indeed appreciated.

REFERENCES

- [1] Algorithmic Inferencing of Aesthetics and Emotion in Natural images : an expansion Ritendra Datta, Jia Li, and James Z. Wang The Pennsylvania State University, University Park, PA 16802, USA.[Online]. Available <http://infolab.stanford.edu/~wangz/project/imsearch/Aesthetics/ECCV06/datta.pdf>
- [2] Optimizing Photo Composition Ligang Liuly Renjie Chen1 Lior Wolf2 Daniel Cohen-Or2 1Zhejiang University, China 2Tel-Aviv University, Israel. [Online]. Available <http://www.math.zju.edu.cn/ligangliu/cagd/projects/composition/>
- [3] Rule of Thirds Detection from Photograph Long Mai, Hoang Le, Yuzhen Niu, and Feng Liu Department of Computer Science Portland State University Portland, OR, USA. [Online]. Available <http://web.cecs.pdx.edu/~fliu/papers/ism2011.pdf>
- [4] Long Mai, Hoang Le, Yuzhen Niu, Yu-Chi Lai and Feng Liu, "Detecting rule of simplicity from Photos", 2011 IEEE International Symposium on Multimedia, 2010
- [5] High Level Describable Attributes for Predicting Aesthetics and Interestingness Sagnik Dhar Vicente Ordonez Tamara L Berg Stony Brook University Stony Brook, NY 11794, USA. [Online]. Available tlberg@cs.stonybrook.edu
- [6] "Rule of Thirds Detection from Photograph", 2011 IEEE International Symposium on Multimedia, 2011, Long Mai, Hoang Le, Yuzhen Niu, and Feng Liu, [online] Available ieeexplore.org
- [7] Frederique Crete, Marina Nicolas, Thierry Dolmiere, and Patricia Ladret, The Blur Effect: Perception and Estimation with a New No-Reference Perceptual Blur Metric paper in SPIE Electronic Imaging Symposium Conf Human Vision and Electronic Imaging, 2007. Available <https://hal.archives-ouvertes.fr/hal-00232709/document>
- [8] Niranjan D. Narvekar and Lina J. Karam, An Improved No-Reference Sharpness Metric Based on the Probability of Blur Detection, in Image Processing, IEEE Transactions on (Volume:20 , Issue: 9), 2011. [online] Available ieeexplore.org
- [9] Long Mai, Hoang Le, Yuzhen Niu, Yu-Chi Lai and Feng Liu, Detecting rule of simplicity from Photos, 2011 IEEE International Symposium on Multimedia, 2010. [online] Available web.cecs.pdx.edu/~fliu/project/simplicity/
- [10] Ligang Liu, Renjie Chen, Lior Wolf, Daniel Cohen, Optimizing Photo Composition, Zhejiang University, China, Tel-Aviv University, Israel, Nov 2010 [online] Available www.math.zju.edu.cn/ligangliu/cagd/projects/.../Composition-TR.pdf
- [11] Shapiro, L. G. & Stockman, G. C: "Computer Vision", page 137, 150. Prentice Hall, 2001
- [12] Mark S. Nixon and Alberto S. Aguado. Feature Extraction and Image Processing. Academic Press, 2008, p