Classifying the Photo Aesthetics using Heuristic Rules of Photography

S. S. Mankikar
Department of Computer Engineering
MAEER's M.I.T. PUNE
Pune, India

Prof. M. V. Phatak
Department of Computer Engineering
MAEER's M.I.T. PUNE
Pune, India

***Abstract*— with the advent of digital cameras there is boom of multimedia data online. People click photos and post in various online community photograph sites. Thus there is need to archive these photos systematically, since millions of photos are added online, everyday. Although much effective content based image retrieval systems have been proposed, the system can be made more effective if the user gets high quality images for his search. Thus, the motive behind the emergence of Computational Aesthetics is to carry out automatic quality assessment and rate the exponentially growing images online with the aesthetic score. This paper aims to classify the aesthetics of the photographic images taking into consideration the various rules of photography and applying them to the object of interest in the image. For object of interest detection we have used Visual saliency technique and for classifying the aesthetics we have used SVM technique. We classify the images as "aesthetically pleasing" or "aesthetically non-pleasing". We report our Results at the end.**

Keywords*— *automatic quality assessment, Computational aesthetics, Computer vision, HCI, Image processing, SVM.

I. AESTHETICS DEFINITION

According to oxford dictionary means [13] "A set of principles underlying the work of a particular artist or artistic movement".

According to Britannica encyclopedia [14] " the philosophical study of beauty and taste. It is closely related to the philosophy of art, which is concerned with the nature of art and the concepts in terms of which individual works of art are interpreted and evaluated."

II. AESTHETICS IN PHOTOGRAPHS

Aesthetics by definition means the ethics or philosophy related to the appreciation of beauty. The word aesthetics is related to art and has been recently related to the field of photography. Aesthetics in images may arouse a pleasing or non –pleasing emotion in a human. These emotions can be quantified as well as classified with the help of machine learning algorithms. This field may sound subjective but following the basic guidelines that professional photographers follow to take pleasing photographs, a system can be built to classify the aesthetics of an image.

With the proliferation of cheap cameras there has been boom in the archiving the digital images online, not only digital images but also multimedia data. Various image retrieval techniques and annotation techniques has been proposed, and incorporated in image search. With growing number of photos shared by people every day, thanks to online community sites such as Flickr [26], Photo.net [27], DPChallenge [14], Instagram [28] and many more. These photos can be classified according to the aesthetics to give users pictures with high quality for their search. In this paper first we discuss the motivation and the applications in this field. Then in section v we discuss the existing systems. Further we discuss our proposed architecture and its execution. Finally we report our results and conclusion at the end.

III. MOTIVATION

Taking into consideration the growing number of photos on photo sharing sites there is a need to classify these photos according to their aesthetic appeal. This kind of systems can be further embedded in many applications, which is described in the forthcoming section.

IV. APPLICATIONS

1. Aesthetics can be taken into consideration for designing appealing flash cards. Appealing flash cards can be used in educational purpose for effective learning of students.

2. Still advertisements are one of the important elements in marketing the product. Catchy and aesthetically pleasing still advertisements grab viewer's attention.

3. Web sites are speedily becoming the desired media choice for data search, online shopping and for educational purpose. In this case aesthetics play an significant role in appealing to the taste of the targeted audience.

4. It can be used to build personalized photo album, where the photos of the album are chosen according to the aesthetic appeal.

Fig 1. Various Applications of Computational Aesthetics

V. LITERATURE SURVEY

We studied various papers in literature and examined that the frameworks built till date followed either a top down approach or bottom up approach.

An early attempt in this field was using a black box approach that could classify photos as "Amateurs" and 'Professionals" [15] they implemented low level features and trained a classification framework (SVM and naive Bayesian). But their work couldn't establish any relation between visual photographic features and low level features.

Apart from establishing relation between low level features of images and high level aesthetics an attempt was made to model metadata of photo in the aesthetic classification system. Metadata is generated automatically in cameras. This automatically generated metadata contains size of the file, date at which photograph was taken. This metadata was incorporated by Boutell et al [10] to classify the indoor outdoor and the sunset scenes. The drawback of the system is, the metadata of all the photos is not available.

The first attempt to build such correlation was by [2] where they established a mapping between the low level features of images to high level features of photography such as rules of third, Low depth of field etc. Their features were weak and resulted in low classification rate.

[3] Determined perceptual factors such as 'simplicity' and 'realism' which are the reasons that distinguish the professional photographs from the "snapshots" and modeled these factors into low level feature. Hence the approach came to be known as the "top down approach".

These ideas were further taken on next level where a relationship between the background and foreground of the photo could be established. Subject in the photo was extracted by various visual attention models and then Rules were applied to it. It is so, because it is believed that subject of the photo grabs the attention of viewer. Thus several attempts [16, 17, 18, 19 and 20] focused on the subject of the photo and extracted features around the subject. While [17] used saliency map approach by [21] for identifying multiple subjects form the photo and k means clustering to extract the subjects from the photo. On the other hand [19] extracted the subject area by their three proposed methods: Clarity based subject area detection, Layout based subject area detection and Human based detection. [16, 18] also emphasized on the subject of the photo.

[22] Followed an instance based approach. They kept their system simple .They kept their framework simple by not applying any subject region detection or image segmentation technique as they established the fact that it increases the complexity and greater time for computation

[8] Proposed high level describable attributes such as "presence of salient objects", 'Portraits and people, "Scene" etc. they improved on the results of [29] which indicated that incorporation of high level attributes lead to improved performance in such systems.

[23] Proposed generic image descriptors for evaluating the aesthetic quality of photos. Their generic image descriptors: Bag of visual words and Fisher vectors performed well on large dataset. They used SVM and Stochastic Gradient Descent algorithm for classifying the features. Their proposed features outperformed various state of the art works' features [2, 19] which used conventional rules of photography.

[24] Proposed relative features by comparing with group of similar photos. They also detected the incompleteness in the photos which may occur when main subject of the photo is excluded.

VI. PROPOSED WORK

Figure 1 depicts the architecture of our system. The data set for this system is acquired from DPChallenge website [14]. To find out the focus of attention in the image we need to calculate the saliency map. After calculating the saliency map, saliency value of the object of interest is computed. Then features are extracted based on the rules of photography and a training set is created. After the training set is created SVM classifiers are trained to classify the photographic image given as input. The detailed implementation is discussed in forthcoming subsections.

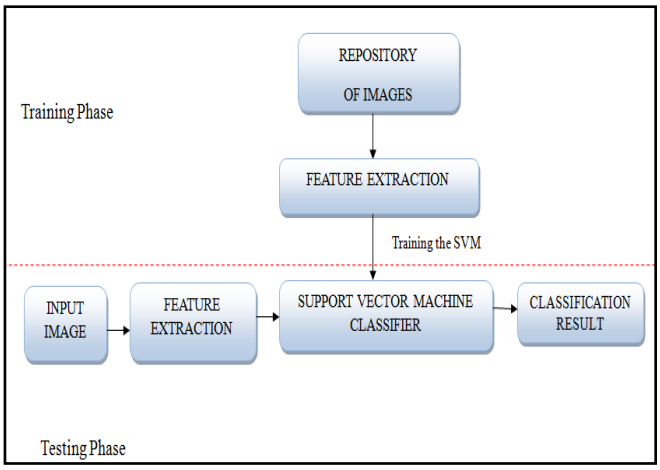

Fig 2. Architecture of proposed system

1) Detecting the object of interest in photographs

In order to find out the subject of the photo, normal object detection algorithms cannot be of great help. It requires visual saliency detection algorithms to determine what catches the human attention. Various visual saliency detection algorithms are proposed till date but out of these we chose Achanta's saliency detection algorithm for finding out the visually salient regions in photographs [7] as it is simple, the saliency map is of the same size as that of image and fast in computation.

Achanta et al's algorithm is based on both local contrast and sliding window technique. The saliency map is calculated at three different scales and combined to get the final saliency map.

2) Rules of Photography as features

Photographers follow certain rules of photography for capturing images. These rules of photography distinguish the professionally shot photos from the ordinary ones. We call these rules of photography as "High level features". The Low level features such as color, texture, shape, edges etc can be modeled to map these high level features or vice versa depending on the approach (top down or bottom up) used to build the system.

The subject of interest captures human attention. Keeping into mind this simple principle some rules of photography has been designed to focus on subject of interest in the photos. Some of these rules of photography are rules of third, depth of field , simplicity etc. In context to image processing and computer vision domain, the first step to implementing these rules of photography is by identifying subject of the photograph. In order to classify the photos we chose four subject driven rules of photography i.e. rules of third, depth of field , diagonal dominance, visual balance and rules of simplicity to apply to the salient object. The rules are described in detailed as follows:

Rules of third: A photo where the main subject is located near one of the four power points formed by the intersection of two equally spaced horizontal and vertical lines is to follow rules of third.

1. Calculate the power points in an image that we get after intersection of horizontal and vertical third lines

2. Calculate the centroid C of the salient region. The centroid of the salient region is calculated by regionprops function
3. Calculate the distance of the centroid from each of the power points.
   $D_m(C, pow_i)$; i= {1, 2, 3, 4}, $D_{m=}$ Manhattan distance

4. If the distance is minimum then the image is said to follow rules of third

Depth of field: The region of interest is in sharp Focus and the background appears blurred [6]**.** Photographers prefer low depth of field for such kind of photography. The depth of field where the background is blurred to highlight the subject, the blur factor in the image can be calculated. To calculate the blur factor we used blur metric indicator by Crete et al. [1].

1. Read the image
2. Apply low pass filter to it
3. Calculate the intensity variations in original image
4. Calculate the intensity variations in blurred image
5. Compare the intensity variations
6. Calculate and normalize the Blur metric.

Rule of Diagonal dominance: Diagonal lines draw the attention of a viewer. A subject aligned diagonally adds aesthetic appeal to the image [5].

We implemented Renjie et al's [6] rule of diagonal dominance feature. Where the distance between the prominent lines and the diagonal lines is. The Prominent lines are identified by the Hough transform algorithm. Once the prominent lines are identified select those lines which support the salient region. Then the distance between the diagonal lines and the prominent lines is calculated.

Visual balance: A picture may be a balanced one in that its major elements are evenly distributed but a viewer may not be able to determine the photographer's intention from this arrangement.

We calculate the Visual balance as the distance between centroid of salient region and image centre. If the distance is less then the image is said to be visually balanced as given in [6]

Rule of Simplicity: If an object is placed against simple background then the image is said to follow rule of Simplicity.

For computing rule of Simplicity we implemented Mai et al's [25] features such as compactness and Background Simplicity.

3) Training the SVM

Support Vector Machines are based on the concept of decision planes that define decision boundaries. A decision plane is one that separates between a set of objects having different class memberships. SVM being one of the powerful binary classifiers, it best suited our classification purpose. Based on the features computed the SVM was trained accordingly. For training purpose we used LibSVM with RBF Kernel.

VII. RESULTS

For creating training set and test set we have acquired photographs from DPChallenge [14]. These photographs belongs to the category of floral, fruits, birds etc. We acquired a dataset of 200 images out of which 100 were used for training and remaining 100 images were used for testing.. We intend to expand our dataset to achieve higher precision and recall rate. We achieved a overall classification rate of 79.16%Taking into account the nature of the problem the results are promising.

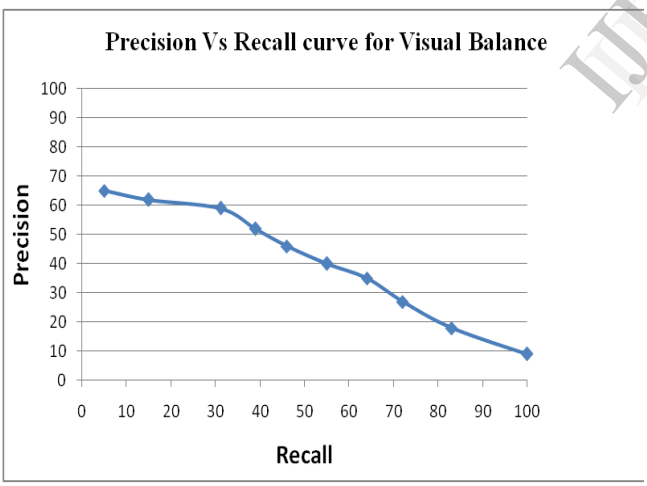

Fig 3. Precision vs Recall Curve of Rule of Visual balance

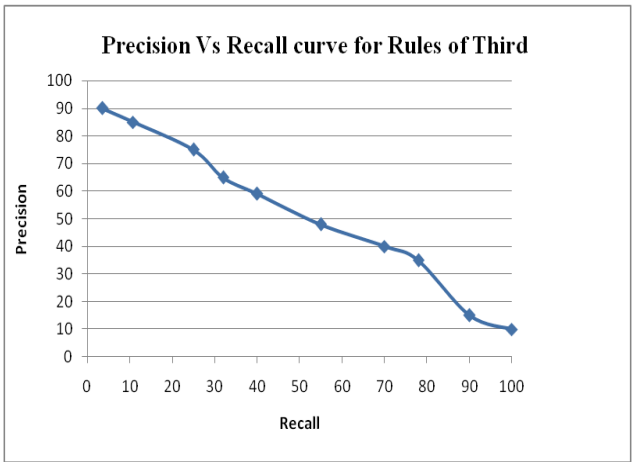

Fig 4. Precision vs Recall Curve of thirds

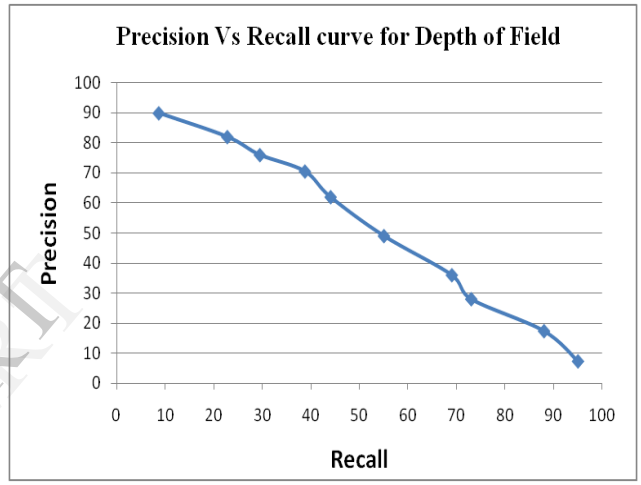

Fig 5. Precision vs. Recall Curve of Depth of Field

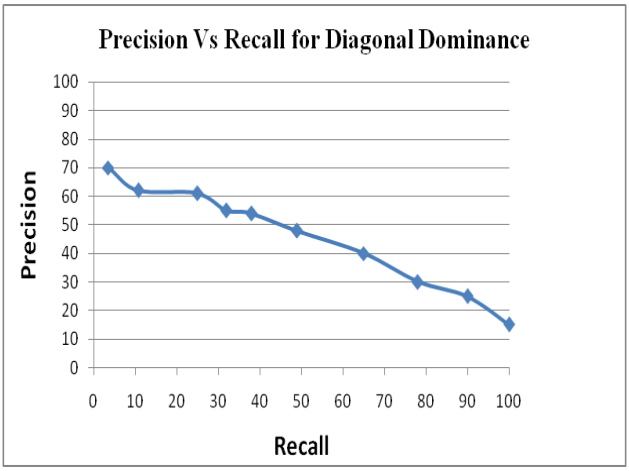

Fig 6. Precision vs Recall Curve of Diagonal dominance

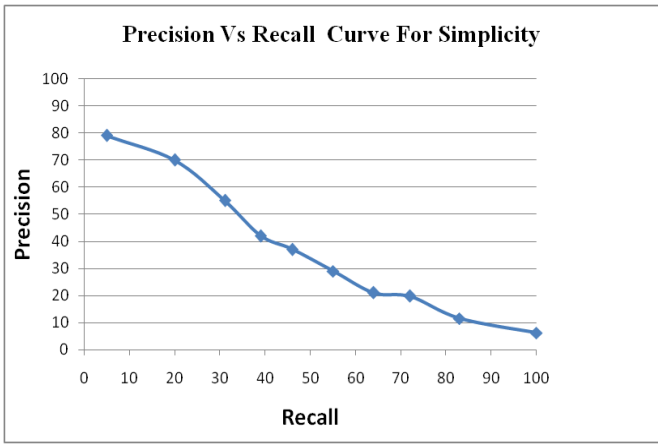

Fig 7. Precision vs recall of Simplicity

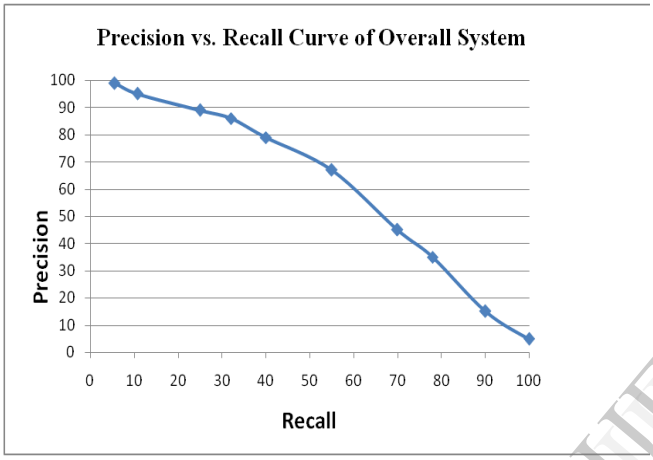

Fig 8. Precision vs. Recall Curve of Overall system.

VIII. CONCLUSION

In this paper, we proposed a system which could classify a photo as "aesthetically pleasing" or "aesthetically non-pleasing" based on the rules of photography features which focus on the subject of the photographs and we achieved a classification accuracy of 79.16%. The first phase is identifying the subject of the photo by visual saliency detection algorithm. The second phase is to compute the features based on the rules of photography. The third phase id training the SVM. We intend to add more rules of photography focusing on subject to improve the classification accuracy. We also intend to find out the aesthetic score of the images in future.

VIII. REFERENCES

[1] F. Crete, T. Dolmiere, P. Ladret, and M. Nicolas, "The blur effect: perception and estimation with a new no-reference perceptual blur metric," in SPIE Conference on Human Vision and Electronic Imaging, 2007.

[2] R. Datta, D. Joshi, J. Li, and J. Z. Wang, "Studying Aesthetics in Photographic Images Using a Computational Approach," Proc. of the European Conference on Computer Vision, Part III, pp. 288-301, Graz, Austria, May 2006.

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

[4] Hsiao-Hang Su, Tse-Wei Chen, Chieh-Chi Kao, Winston H. Hsu, and Shao-Yi Chien, "Preference-Aware View Recommendation System for Scenic Photos Based on Bag-of-Aesthetics-Preserving Features", IEEE transactions on multimedia, vol. 14, no. 3, June 2012.

[5] http://digital-photography-school.com/using-diagonal-lines-in-photography

[6] http://www.digital-photography-tips.net/depth-of-field.html

[7] R. Achanta, F. Estrada, P. Wills, S.S ustrunk, "Salient region detection and segmentation", in: Proceedings of the International Conference on Computer Vision Systems, 2008, pp.6675.

[8] Sagnik Dhar , Vicente Ordonez and Tamura L Berg, "High Level Describable Attributes for Predicting Aesthetics and Interestingness", In Proc.IEEE Int'l Conf Computer Vision and Pattern Recognition, 2011.

[9] Ligang Liu Renjie, Chen Lior Wolf and Daniel Cohen-Or, "Optimizing photo composition", Computer Graphic Forum Proceedings of Eurographics volume 29 pages 469-478, 2010

[10] M. Boutell and J. Luo. Bayesian fusion of camera metadata cues in semantic scene classification. In CVPR, 2004.

[11] S.Bhattacharya, R.Sukthankar, and M. Shah. A Framework for Photo-Quality Assessment and Enhancement based on Visual Aesthetics. In *Proc. ACM MM*, 2010.

[12] http://www.DPChallenge.com/

[13] http://www.oxforddictionaries.com/definition/english/aesthetic

[14] http://www.britannica.com/EBchecked/topic/7484/aesthetics

[15] H. Tong, Li Mingjing, Zhang Hong-jiang, He Jingrui, and Zhang Changshui. Classification of digital photos taken by photographers or home users. In Proceedings of Paciffic Rim Conference on Multimedia, pages 198_205. Springer, 2004.

[16] Domen Pogancnik, Robert Ravnik, Narvika Bovcon, Franc Solina," Evaluating Photo Aesthetics Using Machine Learning", Conference on Data Mining and Data Warehouses 15th International Multiconference on Information Society.

[17] M. Nishiyama, T. Okabe, Y. Sato, and I. Sato. Sensation-based Photo Cropping. In Proc. ACM Multimedia, 2009.

[18] S.Bhattacharya, R.Sukthankar, and M. Shah. A Framework for Photo-Quality Assessment and Enhancement based on Visual Aesthetics. In *Proc. ACM MM*, 2010.

[19] Xiaoou Tang , Wei Luo and Xiaogang Wang, "Content based Photo Quality Assessment", in Proc. IEEE transactions on multimedia 2013

[20] Y. Luo and X. Tang, "Photo and video quality evaluation: Focusing on the subject", In ECCV, 2008.

[21] Itti, L., Koch, C. and Niebur, E. A Model of Saliency-Based Visual Attention for Rapid Scene Analysis. IEEE Transactions on Pattern Analysis and Machine Intelligence, 20(11): 1254-1259, 1998.

[22] Kuo-Yen Lo, Keng-Hao Liu and Chu-Song Chen. "Assessment of Photo Aesthetics with Efficiency", International Conference of Pattern Recognition (ICPR2012).

[23] Luca Marchesotti, Florent Perronnin, Diane Larlus, Gabriela Csurka,"Assessing the photo quality of photographs using generic image descriptors", Computer 2011 pp 1784 – 1791

[24] M. Yeh and Y. Cheng,"Relative features for photo quality assessment", In Proc. IEEE Intl Conf. Image Processing, 2012.

[25] Long Mai, Hoang Le, Yuzhen Niu, Yu-chi Lai, and Feng Liu. Detecting Rule of Simplicity from Photos. **ACM Multimedia 2012**, Nara, Japan, October 2012.

[26] http://www.flickr.com/

[27] http://photo.net/

[28] http://www.instagram.com

.