JSMO
DYPCOE

JOURNAL OF SIMULATION MODELING & OPTIMIZATION

Available online at www.jsmo.co.in

Image Enhancement By Discovering Aesthetic Community using Latent Semantic Analysis

Priyanka R. Chaudhary

Department of Computer Engineering D.Y.Patil College of Engineering, Akurdi, Pune.

A B S T R A C T

Image enhancement has become one of the very important techniques in image processing to improve the visual appeal of an image. Image enhancement provide better result in by transforming representation for future automated image processing. The use of image enhancement for image sharing has increased considerably in recent years due to social media like Facebook where every user share their images of any occasion. Nowadays, images can be enhanced by various applications such as Instagram, Microsoft Office Picture Manager, Adobe Photoshop etc. as images are the source of information for interpreting and analyzing data. Image enhancement is the method through which image can be altered with respect to its pixels values. The different four phases of the system are preprocessing, feature extraction, discovery of aesthetic community and image enhancement. Discovering aesthetic community is important to explore number of clusters of aesthetic features from images so that in image enhancement these features can be used. The proposed system uses the latent semantic analysis (LSA) for extracting similar topic to form the aesthetic community. The proposed system works more efficiently as discovery of aesthetic community and image enhancement helps to increase the quality of image. Thus the implementation of image enhancement using discovery of aesthetic communities provides a more efficient of enhancement process.

Keywords: Aesthetic community, community detection, Hue, Image Enhancement, LSA, Topic model

1. Introduction

Image enhancement has become an essential technique in various image applications. It is widely used in printing industry, cinematography, graphic design, medical forensic purpose, image analysis from satellite, atmospheric sciences, oceanography etc. Image enhancement is the process of increasing the aesthetic visual appearance of an image such as resolution change, spatial recomposition, image aspect ratio etc [1]. Due to the digital photography, anyone can take photo as camera is available even in cell phone in cheapest rate. It provides the facility to share the image on social media like Instagram, Facebook etc. as day by day popularity of image sharing sites are increasing. Large amount of images are being captured and shared on social media, due to growing interest in aesthetics quality improvement. The aspect of photography that contributes to high quality images is image composition through its aesthetics features. Image enhancement provides unified way to design aesthetically pleasing images like wallpaper [2]. With this growing interest in aesthetics quality improvement, image enhancement received considerable research interest in the recent years. Most of the researchers have proposed numerous image enhancement techniques that utilizes recomposition, retargeting, colorization to feature in the resulting images. Image enhancement method can be categorized into spatial domain and frequency domain [1]. Spatial domain method works directly on the image pixels. The pixels values are calculated and manipulated to get the desired enhancement. In frequency domain, image is first converted into frequency domain i.e. the fourier transform is used. The enhancement operations are carried out on the Fourier transform of the image and then the Inverse Fourier transform is performed to get the resultant image [1]. The goal of the image enhancement is to improve the understandability or approach of information in images for observers and to provide better input for other automated image processing techniques. A. Problems of Conventional Approaches Many image enhancement algorithms have been proposed in the literature which achieved promising results. Even though after many research the conventional approaches faces many problems which are as follows: • An ideal photo enhancement framework can be build by professional photographers with their experiences but it will be difficult for the photographers to give each aesthetic topic and add their experiences into image enhancement system. • Describing the aesthetic interest of a photographers is difficult task as most of the photographers upload thounsands of the images from image hosting site like Flicker while some peoples upload very few images, it will cause the overfitting problem. • Image tags reflect human visual perception at high level but along with image appearances to model the aesthetic

* *Corresponding author.* Tel.7410158595.
E-mail address: *priyanka.sherekar@yahoo.co.in*

interest is challenging as many image tags are noisy or missing. • As image enhancements includes a series of operations like retargeting, recompositions, colorization but few operations can guarantee that the aesthetic cues are optimally preserved. • Many image enhancement algorithms rely on human computer interaction to achieve good results. This requirent might be impractical for large scale image processing. To solve the above problem, a unified framework is proposed to enhance the image by detecting aesthetic communities using topic model. For making framework various image enhancement tasks are combined using a probabilistic model. For regularized topic model firstly have to extract the tags associated with images. Most of the images uploaded on image hosting site like Flickr have tags associated with it. Tags are the extra information with images. Aesthetic community defines the users of having similar aesthetic topics which are densely connected. The image enhancement framework defines the socially aware model that gives more image attractiveness which can be increased due to detecting aesthetic communities. A classic topic model with aesthetic topics of Flickr users describes the latent topics. The unified image enhancement framework describes the visual features which helps in enhancing the given image [2]. Many image aesthetic models and community detection approaches have been addressed well in literature. Section I defines introduction about image enhancement and its applications, section II includes Literature Review about aesthetic model and community discovery approaches and section III includes Implementation Details, section IV describes Result analysis and section V includes Conclusion.

2. Materials & Methods

Image enhancement is the subarea in image processing has become a popular research topic because of increased social awareness. Enhanced images increases the visual attractiveness. The image aesthetic enhancement algorithm covers the image analysis and data mining domain. Unified image enhancement framework which can be used to detect the aesthetic topic and forming the different communities using similarity measure between them. The related work is divided into two parts: image aesthetic models and community detection are as follows. A. Image Aesthetic Models Aesthetic image means artistic, creative image including its beauty. The aesthetic specifies the critical reflection on art, culture and nature. Aesthetic also refers to a set of principles underlying the works of a particular art movement or theory. Image aesthetic model gives the visual features extracting from the image. There are various conventional aesthetic models are researched in present days. Anu Joy et al. explains how the aesthetic is a special session in the current area of computational visual aesthetic [3]. They mainly focuses on survey of aesthetic quality classification of photographers including the image quality factors, its assessment and different approaches [3]. R. Datta et al. proposes a significant correlation between various properties of photographic images and their aesthetics ratings. They used a community based database and ratings that gives visual properties for aesthetic quality and the global image features [4]. Wong et al. extracted three types of visual features which includes global image features, salient regions features and features that depict the subject background relationship [5]. Lind [6] explains aesthetic objects are important and intresting to attract the attention of the user. W. Luo et al. proposes content based photo quality assessment which mainly focuses on subject area detection with new global and regional features [7]. For showing features, author uses the hue wheels and prominent lines mainly. Object is essential for detecting the image. In [8] author presented an objectness measure which is used to differentiatethe object from the background ones of an image. Dhar et al. explains the high level describable attributes (e.g. Mountain, sunset) which measures gives information of images i.e. image composition, its content and sky-illumination attributes [9]. Zang et al. proposed the fusion of visual features, local and global structural cues for photo aesthetics [10]. Wang et al. proposes human scenery positional relationship in which main focuse is on human subjects in a background scene [11]. An android based photographic assistance system is developed which produces well-composed souvenier photos. B. Community Detection Community is a group that have a particular characteristics in common. It indicates the situation of sharing or having certain attitudes and interests in common. The goal of community detection is to identify the structure of community in network. A community also defines clusters of densely connected vertices that are loosely connected to other vertices. Detecting communities in networks is a big challenge, Lanchichinetti et al. introduces a new class of graphs for testing algorithms which is used to identify communities [12] by testing modularity optimization and a clustering technique based on Pott's model. Jure Leskovec et al. [13] gives the range of network community detection methods and results shows that determining clustering structure of large networks is very complicated. Yoshida proposes profile graph i.e. graph structure constructed from the profile data which builds a simple model for connectivity and profile graph over a social network data [14]. Yang et al. finds that the community overlaps are more densely connected than the non-overlapping parts [15]. Author proposes community affliation graph model which is a model based community detection method that constructs on bipartite node community affiliation networks. S. Gregory introduces a new algorithm for detecting overlapping communities and describes the CONGO algorithm (Cluster Overlap Newmann Girvan Algorithm Optimized) which overcomes the problem of CONGA algorithm by making it much faster using local betweenness [16]. Y. Zang et al. tries to detect community structures from large network data proposes a classic dyanimic community detection algorithm uses propinquity which is a measure of the probability for a pair of nodes involved in a coherent community structure [17]. Latent

Semantic Analysis (LSA) [18] is a topic extraction model in a set of corpus of documents (Flickr users). It is used to retrive the user and gives the similarities between the images of the users. Hao Zang et al. proposes the RG-PLSA model (Regularized Guassian Probabilstic Latent Semantic Analysis) that is a hybrid verion of Guassian PLSA's and Hierarchical Guassian mixture models(HGMM) which tries to learn aesthetic communities and try to prevent overfitting and provides flexibilty [19]. LDA(Latent Dirichlet Allocation) is a topic model [20] along with its different forms [21] used to discover the communities which are socially aware. It uses serial communities users to form the social link graph and differentiate them into various communities. Probabilistic models are also used for this purpose. Social context are important for finding the multiple people as object through images, in [22] author proposes the socially aware model to find many people occupations through visual attributes. Even for finding the communities the email-corpus is used in [23] which extracts the e-communities from email corpus showing main part as how there is social interactions and topic matching mining from documents to search communities. Yin. et al proposes a text associated graphical model which gathers the generation of links between users and words of users to extract communities giving the output as link structure with similar semantic topics [24].

3. Implementation Details

The main purpose of proposed system is to enhance the image by using tags of the images with high accuracy. For enhancing the image the features associated with it are gathered. Aesthetic communities are formed whose having the similar topics of each image users. On the basis of it, different densely connected aesthetic communities are formed. Finally the maximum features are transferred by using probability model. A. System Architecture Fig. 1 shows the architecture of the proposed system which comprises taking input image from dataset. System architecture demonstrate the fully automatic framework which is proposed to enhance and increase the image visual appearence by discovering communities consists of aesthetics by using a latent semantic analysis topic model. Probabilistic model combines different image enhancement techniques. Firstly, extract a set of visual features for expressing each image from multiple channels. A tagwise regularizer latent semantic analysis (LSA), a topic model is used which is learned from image appearence features and the aesthetic topic of each Flickr user is shown by latent aesthetic topics. KL divergence is used to represent the similarity of aesthetic topics among users. These users having similar aesthetic topics are closely related. These are used to form the aesthetic communities. Finally in the probabilistic model, the extracted visual

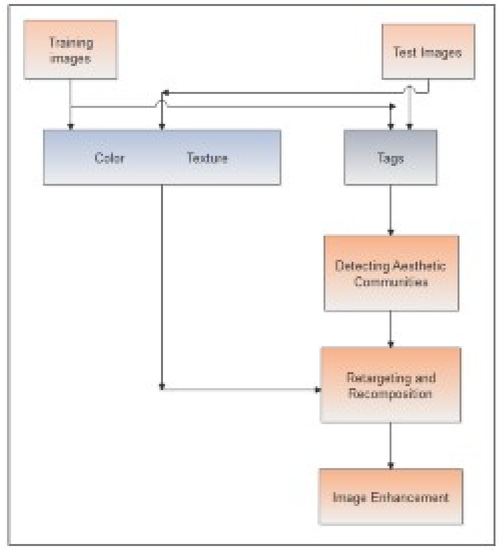

Fig 1. System Architecture

features i.e. color and texture are transfer to the test image and various image enhancement operations mainly retargeting and recomposition is used to enhance the image. This system architecture is used to build the socially aware model which helps to grow the image attractiveness. This can be achieved by discovering aesthetic communities from images by discovering topic model. The visual features or appearence is used to make the image enhancement framework.

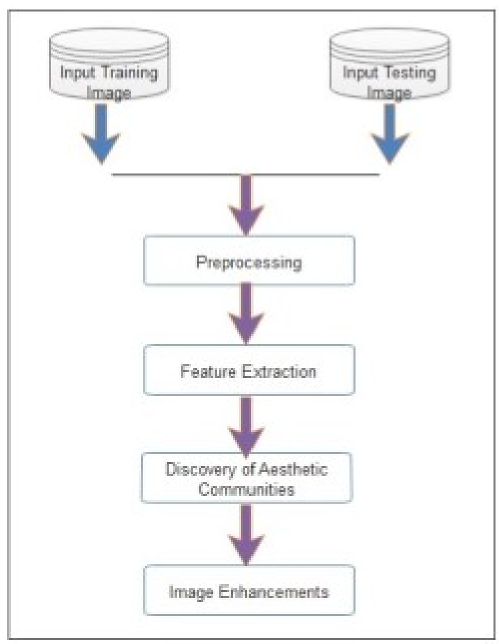

Fig 2. System Overview

The proposed system overview consists of the following steps as shown in Fig. 2 Preprocessing: In this phase, the original image is taken from the data set then color and texture are used to obtaineimage aesthetic features which gives the hue, satuaration, value and canny edge image. •

Color Channel Colors are the representation of the image. It is represented in the form of harmonic template. Harmonic template is the property that contain aethetically pleasing color according to human viewpoint. These colors are represented by hue wheels are as follows:

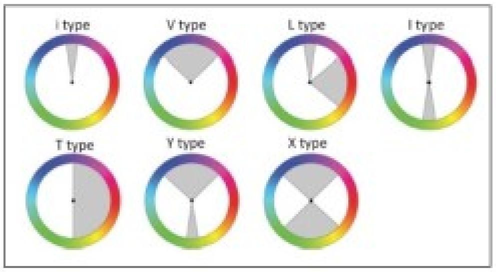

Fig 3. Seven hue wheels in harmonic template

Hue wheels gives either single sector or double sector and it indicates the gray color in which colors are fallen are fixed. For finding color harmony of each image, firstly see whether it will comes under single sector T1 or double sector T2.

$$\arg\min_{T_k} \sum_{i \in I} D(H(i), T_k).S(i) + \alpha A(T_k) \qquad (1)$$

where H(i) and S(i) denote the hue and saturation of the ith pixel in photo I respectively; $D(H(i),T_k)$ is the arclength of H(i) to the closet sector border and if H(i) falls into the sector of the template, then D(H(i), Tk) = 0. Large arc-length sectors are mostly fitted by an image, therefore setup a regularized term where $A(T_k)$ is the arc-length (single-sector) or accumulated arc-length (double-sector) of each harmonic template and α is a free parameter. It is checked whether image I can be fitted by a singlesector or double-sector harmonic template, characterize the color distribution of an image by a four-dimensional feature vector:

$$x_c = \begin{cases} (\alpha_1, s_1, 0, 0) & if \;\; I \;\; in \;\; T1 \\ (\alpha_1, s_1, \alpha_2, s_2) & if \;\; I \;\; in \;\; T2 \end{cases} \qquad (2)$$

where α1 and α2 denote the center angles of the sectors. s1 and s2 are the average saturation of harmonic colors, i.e., the colors inside the gray areas of a hue wheel. Adopt color saturations here because the low and highly saturated harmonic colors are distinguishable based on human perception. • Texture Channel For finding the textural property of an image the prominent lines are used. The location and gradients of prominent lines are utilized for this channel. It is extracted by Hough transform and then differentiate into horizontal and vertical lines according to slope. A horizontal prominent line having slope between -1 and +1 and vertical if its slope is larger than +1 or smaller than -1. The average orientations of horizontal and vertical lines are denoted as o¯h and o¯v ; the average vertical position of horizontal lines as p¯ v h and the average horizontal position of vertical lines as p¯ h v , a four-dimensional feature vector:

$$x_t = (\bar{o}_h, \bar{o}_v, \bar{p}_h^v, \bar{p}_v^h) \qquad (3)$$

Represent the appearance of each image by a 8-dimensional feature vector x = $[x_c, x_t]$.Tag Channel Tags reflect the information about image. It explains the semantic of an image associated with it. Like color and texture, the tags assist to learn the latent aesthetic topics from images. Creating bag of words extract the most frequent tags from the training set. Feature Extraction: After detecting color and texture features, hough line transform is used to show prominent line indicated in houghline image. Discovery of aesthetic communities using LSA: Latent Semantic Analysis is a topic extraction model likewise LDA which is used to extract the topic from corpus of users. The similar aesthetic values share similar latent aesthetic topics. A similarity measure is used to describe aesthetic similarity between users. Thses similar aesthetic topics of users belong to the same community so forming the aesthetic communities. Image Enhancement: Image Enhancement can be represented as transferring maximum aesthetic features from training images to target images from various aethetic communities. Firstly retargeting is used in which a greedy scheme is utilized in which image is divided evenly into a number of grids then sequentially shrink either each row or column. After retargeting, recomposition is used in which the objectness measure determines different object regions in an image and move it one by one to an optimal position in the target image.

B. Algorithm The graph shifting algorithm is iteratively used to get the aesthetic community where for each community 5-component Gaussian Mixture Model (GMM) is utilized in image such as:

$$p(x \mid \theta) = \sum_{i=1}^{5} \eta_i N(x \mid \varphi_i, \Sigma_i) \qquad (4)$$

where x is the 8-dimension color and texture feature vector of an image, {η, ϕ, Σ} are the GMM parameters. Every aesthetic community contains Flickr users with a specific aesthetic topic. The probabilistic model combines connection between training image with test image. The five layers are represented. The first layer shows the training images I 1 , ..., IL, the second layer shows the visual features X from the training imagess, the third layer is the detected aesthetic communities C, the fourth layer represents the visual feature x from a test photo, and the last layer denotes a test photo I .

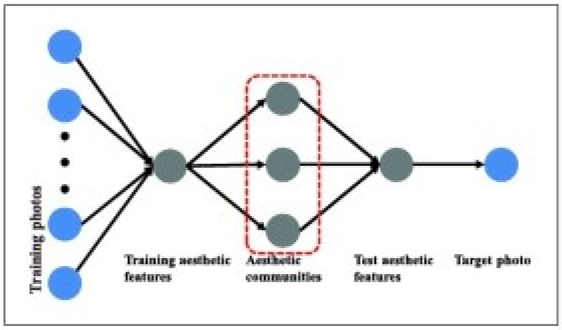

Fig. 4. Probabilistic model for image enhancement [2]

Image enhancement can be represented as maximally transferring aesthetic features from the training photos into a test photo, based on the perception of Flickr users from different aesthetic communities. This process can be formulated as:

$$\max_{I^*} p(I^* \mid I^1, ..., I^L) = \max_{I^*} p(I^* \mid x^*).p(x^* \mid C). p(C \mid X).p(X \mid I^1, ..., I^L) \quad (5)$$

where probabilities $p(I^* \mid x^*), p(C \mid X), p(X \mid I^1, ..., I^L)$ are all equal to one. The probability p(x ∗ | C) is calculated are as follows:

$$p(x^* \mid C) = \prod_{i=1}^{H} p(x^* \mid C_i) = \prod_{i=1}^{H} p(x^* \mid \theta_i) \quad (6)$$

where H is the number of detected aesthetic communities. Ci denotes the i th aesthetic community and θi is the corresponding GMM parameters asshown in equation 6.1. Equation 5 does not have an analytic solution as all the pixels within a image have to be opitimized simultaneously. A gridy scheme is used to solve this problem. Divide a photo evenly into a number of grids for retargeting, then sequentially shrink each column (or row) until p(I ∗ | I 1 , ..., I^L) is maximized. For recomposition use the objectness measure to localize different object regions in a photo and index them based on the posterior probability p(obj | region) then move each of these regions one-by-one to an optimal position in the test photo, aiming to maximize p(I ∗ | I 1 , ..., I^L) Finally, render the background area to make the resulting photo look natural. C. Mathematical Model The mathematical model is a system representation using mathematical functions. The mathematical model for image enhancement using Latent Semantic Analysis (LSA) are asfollows: Let S be the system represented as:

$$S = \{I, F, O\} \quad (7)$$

Set S System for image enhancement. • {I = ij |ij is set of all Images in dataset} {F = Fj |Fj is set of modules used} Here, Fj = {f1, f2, f3, f4} where, f1= Preprocessing of input image, f2= Feature Extraction, f3= Aesthetic community detection and f4= Image enahnacement – f1= Preprocessing of input image. In this multichannel descriptor is used. i.e. color and texture. Color channel is calculated mainly from harmonic template. The color harmony of each image is quantified by deciding whether it should be fitted by a single-sector template T1 or a double-sector template T2:

$$\arg\min_{T_k} \sum_{i \in I} D(H(i), T_k).S(i) + \alpha A(T_k) \quad (8)$$

where H(i) and S(i) denote the hue and saturation of the ith pixel in photo I respectively; D(H(i), Tk) is the arc-length of H(i) to the closet sector border, and if H(i) falls into the sector of the template, then D(H(i), T^k) = 0. It gives the centre angles of the sectors and average saturation of harmonic colors. From texture edges are detected. f2= Feature Extraction. In this prominent lines are detected by Hough line transforms. The average orientations of horizontal and vertical lines are denoted as oh and ov ; the average vertical position of horizontal lines as p and the average horizontal position of vertical lines as p , a four-dimensional feature vector:

$$x_t = (\bar{o}_h, \bar{o}_v, \bar{p}_h^v, \bar{p}_v^h) \quad (9)$$

f3= Aesthetic community detection. For discovering aesthetic community latent semantic analysis is used where topics are extracted and similar topics forms the community. It is generally written as,

$$X = U \Sigma V^T \quad (10)$$

f4= Image enahnacement The detected communities capture the aesthetic topics of multiple Flickr users. They are integrated into a probabilistic model to tackle different image enhancement tasks i.e. retargeting and recomposition.

O1 = O1j |O1j is a Retarget image • O2 = O2j |O2j is a recomposition image O1 and O2 represents the output of the system. Finally the parameter metrics i.e BER can be used for high accuracy. Formula of which is provided in result section.

D. Software Requirement Specification • Hardware Configuration – Processor: Pentium IV or above – RAM : 2 GB(min) – Hard Disk: 20 GB(min) – Key Board: Standard Windows Keyboard – Mouse: Two or Three Button Mouse – Monitor: SVGA

Software Configuration – Operating System : Windows XP or above – Programming Language : JAVA(JDK 1.7), OpenCV – Front End : JavaCV – IDE : Eclipse or Netbeans for Java

4. Experimental Result

A. Dataset The data set is generated by collecting different images from sites like Flickr, Instagram that are freely available online. Images are stored in different public groups like architecture, nightlights, blackwhite, orangeblue, tree , windowseat etc. Each community contains 16 images. The mask images are also present. Tag dataset are created in which differnt tags are written from images. Test tag dataset are also created which contains information about every image. B. Results The parameter metrices such as accuracy, precision, recall, BER is used to evaluate performance of the system. Results of the aesthetic community discovery through BER shows the high accuracy. The balanced error rate is the average of the errors on each class, its defination are as follows:

$$BER = \frac{1}{2}\left(\frac{\#positiveinstancespredictedwrong}{\#positiveinstances} + \frac{\#negativeinstancespredictedwrong}{\#negativeinstances}\right) \quad (11)$$

The optimal match is calculated by the via the hungrarian algorithm [25] as follows:In the Equation 12, f is correspondence between predicted community C and ground truth C¯, dom(f) is set of all possible f values.If the number of predicted communities is less than the number of ground truth then community C has a match C*.

$$\max_{f:C\to C} \frac{1}{|f|} \sum_{C\in dom(f)} (1 - BER(C, f(C))) \quad (12)$$

TABLE I
BER SCORES OF DIFFERENT COMMUNITIES

Groups	BER	Optimal Match
Fabric	0.21	0.78
Glass	0.11	0.88
Leather	0.23	0.76
Metal	0.35	0.65
Average	0.22	0.76

Table 1 shows groups which contains the aesthetically pleasing images. To calculate the adjustment between a predicted community and a ground truth group, Balanced Error Rate (BER) is calculated. Many time the similarity is unknown between detected aesthetic communities and the ground truth group so the optimal match is explained.The proposed system for image enhancement using LSA is under implementation phase. While image collection i.e dataset, preprocessing, feature extraction phases are implemented. The existing system gives BER rate as 0.76 with optimal match.Proposed approach focuses on improving BER rate by 5-10 % than existing approach. Fig. 5 shows the graph of BER comparison of existing algorithm and proposed algorithm, where x-axis represents algorithms and y-axis represents BER score(in percentage).The exactly mining aesthetic communities can enormously improve image enhancement.

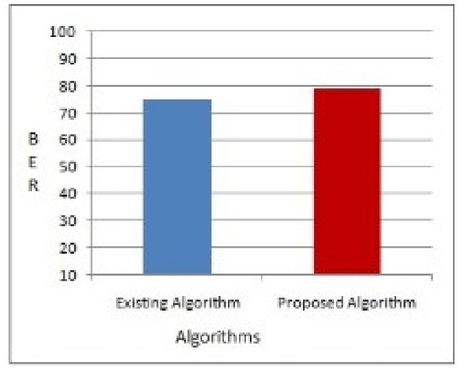

Fig. 5. Comparison of average accuracy of existing and proposed approach

5. Conclusion and Future Work

Social media provides the prevelence of sharing images therefore the image enhancement becomes very important and also challenging with respect to users perspectives. Learning aesthetic communities from a large number of users, where each community contains users with similar aesthetic topics. Since number of images have no tags and some users may upload very few images. Hence it is most important to effi- ciently extract visual features and coherent, interpretable topics to form the aesthetic communities which helps to enhance the image. In existing system aesthetic communities are properly detected but it will be difficult to detect the respective aesthetic community which explains abstract concept (e.g. Jump project). Abstract concept may not gives the proper topic selection expectation. So, the proposed system will try to overcome such problem with more accuracy and less complexity than existing approach. The dataset is created with proper images and all the cases can be tested properly. Results shows the error rate of different aesthetic communities which include four groups i.e. Fabric, Glass, Leather and Metal. By increasing the number of groups can improve the result. In Future work, extend this framework to support more image enhancement operations e.g. colorization and will try to find out that how much the system is accurate by parameters like precision, recall, f-measure, sensitivity and specificity.

Acknowledgements

The authors would like to thank the researcher as well as publishers for making their resources available and the teachers for their guidance. We also thank the college authorities for providing the required infrastructure and support. Finally, we would like to extend a heartfelt gratitude to all friends and family members.

REFERENCES

[1]. Mudigonda, Shanmukha Priya, and Koustubha Priya Mudigonda. "Applications of Image Enhancement TechniquesAn Overview." MIT International Journal of Computer Science and Information Technology, Vol. 5, No. 1, January 2015, pp. 17-21

[2]. Hong, Richang, Luming Zhang, and Dacheng Tao. "Unified photo enhancement by discovering aesthetic communities from flickr." IEEE Transactions on Image Processing 25.3 (2016): 1124-1135.

[3]. Joy, Anu, and K. Sreekumar. "Aesthetic Quality Classification of Photographs: A Literature Survey." International Journal of Computer Applications 108.15 (2014).

[4]. R. Datta, D. Joshi, J. Li, and J. Z. Wang, "Studying aesthetics in photographic images using a computational approach," in Proc. Eur. Conf. Comput. Vis., 2006, pp. 288301.

[5]. L.-K. Wong and K.-L. Low, "Saliency-enhanced image aesthetics class prediction," in Proc. 16th IEEE Int. Conf. Image Process., Nov. 2009, pp. 9971000.

[6]. Lind, Richard W. "Attention and the aesthetic object." The journal of aesthetics and art criticism 39.2 (1980): 131-142.

[7]. W. Luo, X. Wang, and X. Tang, "Content-based photo quality assessment," in Proc. IEEE Int. Conf. Comput. Vis., Nov. 2011, pp. 22062213.

[8]. B. Alexe, T. Deselaers, and V. Ferrari, "What is an object?" in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Jun. 2010, pp. 7380.

[9]. S. Dhar, V. Ordonez, and T. L. Berg, "High level describable attributes for predicting aesthetics and interestingness," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Jun. 2011, pp. 16571664.

[10]. L. Zhang, Y. Gao, R. Zimmermann, Q. Tian, and X. Li, "Fusion of multichannel local and global structural cues for photo aesthetics evaluation," IEEE Trans. Image Process., vol. 23, no. 3, pp. 14191429, Mar. 2014.

[11]. Y. Wang et al., "Where2Stand: A human position recommendation system for Souvenir photography," ACM Trans. Intell. Syst. Technol., vol. 7, no. 1, 2015, Art. ID 9.

[12]. A. Lancichinetti, S. Fortunato, and F. Radicchi, "Benchmark graphs for testing community detection algorithms," Phys. Rev. E, vol. 78, p. 046110, Oct. 2008.

[13]. J. Leskovec, K. J. Lang, and M. Mahoney, "Empirical comparison of algorithms for network community detection," in Proc. 19th Int. Conf. World Wide Web, 2010, pp. 631640.

[14]. T. Yoshida, "Toward finding hidden communities based on user profile," in Proc. IEEE Int. Conf. Data Mining Workshops, Dec. 2010, pp. 380387.

[15]. J. Yang and J. Leskovec, "Community-affiliation graph model for overlapping network community detection," in Proc. IEEE 12th Int. Conf. Data Mining, Dec. 2012, pp. 11701175.

[16]. S. Gregory, "A fast algorithm to find overlapping communities in networks," in Machine Learning and Knowledge Discovery in Databases, vol. 5211. Berlin, Germany: Springer, 2008, pp. 408423.

[17]. Y. Zhang, J.Wang, Y. Wang, and L. Zhou, "Parallel community detection on large networks with propinquity dynamics," in Proc. 15th ACM SIGKDD, 2009, pp. 9971006.

[18]. Anaya, L. H. (2011), "Comparing latent dirichlet allocation and latent semantic analysis as classifiers (Doctoral dissertation)" Retrieved from ProQuest Dissertations and Theses database (University of North Texas, No. 3529209).

[19]. H. Zhang, R. Edwards, and L. Parker, "Regularized probabilistic latent semantic analysis with continuous observations," in Proc. 11th ICMLA, Dec. 2012, pp. 560563.

[20]. D. M. Blei, A. Y. Ng, and M. I. Jordan, "Latent Dirichlet allocation," J. Mach. Learn. Res., vol. 3, pp. 9931022, Mar. 2003.

[21]. G. Costa and R. Ortale, "A Bayesian hierarchical approach for exploratory analysis of communities and roles in social networks," in Proc. IEEE/ACM Int. Conf. ASONAM, Aug. 2012, pp. 194201.

[22]. Shao, Ming, Liangyue Li, and Yun Fu. "What do you do? occupation recognition in a photo via social context," in Proceedings of the IEEE International Conference on Computer Vision, pp. 3631-3638. 2013.

[23]. D. Zhou, E. Manavoglu, J. Li, C. L. Giles, and H. Zha, "Probabilistic models for discovering e-communities," in Proc. 15th Int. Conf. World Wide Web, 2006, pp. 173182.

[24]. Z. Yin, L. Cao, Q. Gu, and J. Han, "Latent community topic analysis: Integration of community discovery with topic modeling," ACM Trans. Intell. Syst. Technol., vol. 3, no. 4, 2012, Art. ID 63.

[25]. Kuhn, Harold W. "Variants of the Hungarian method for assignment problems." Naval Research Logistics Quarterly 3.4 (1956): 253-258.