Discovering Harmony: A Hierarchical Colour Harmony Model for Aesthetics Assessment

Peng Lu[1], Zhijie Kuang[1], Xujun Peng[2], and Ruifan Li[1]

[1] Beijing University of Posts and Telecommunications, Beijing, 100876, China
[2] Raytheon BBN Technologies, Cambridge, MA, USA

Abstract. Color harmony is an important factor for image aesthetics assessment. Although plenty of color harmony theories are proposed by artists and scientists, there is little firm consensus and ambiguous definition amongst them, or even contradictory between them, which causes the existing theories infeasible for image aesthetics assessment. In order to overcome the problem of conventional color harmony theories, in this paper, we propose a hierarchical unsupervised learning approach to learn the compatible color combinations from large dataset. By using this generative color harmony model, we attempt to uncover the underlying principles that generate pleasing color combinations based on natural images. The main advantage of our method is that no prior empirical knowledge of image aesthetics, color harmony or arts is needed to complete the task of color harmony assessment. The experimental results on the public dataset show that our method outperforms the conventional rule based image aesthetics assessment approach.

1 Introduction

With the pervasive use of digital cameras and cell-phones and the deluge of online multimedia sharing communities, image retrieval has drawn much attention in recent years. As revealed in [1] and [2], besides semantic relevance, users prefer more appealing photos retrieved by image search engines, which indicates aesthetics and quality properties of images play a more important role than semantic relevance. Also, by integrating photograph aesthetics models into hand-held devices, a real-time recommendation can be made to facilitate professional and amateur users to manipulate the captured images. Thus, to assess photographs quality automatically from the perspective of visual aesthetics is of great interest in the research of web image search [1] and multimedia processing [3, 4].

Although there have been many researches focusing on the image aesthetics estimation, assessing aesthetics quality of photographs is still an extremely challenging problem because it is a subjective task and different people may have various tastes to the same image. Despite lack of firm consensus and ambiguous definition of aesthetics, there exists several simple criteria to distinguish "good" images from "bad" ones. For example, as shown in Fig. 1, most people agree that the sample photos on the bottom row have higher aesthetics attributes than the images on the top row, which have degradations due to out-of-focus blur, compression artifacts, distortion, etc. Moreover, for professional photographers, various principles (*e.g.* composition, sharpness, proper contrast and lighting as well as special photographic techniques as shown in Fig. 1(d), 1(e) and 1(f)) are taken into consideration to make a photo more attractive. Then, it is possible to design computational methods to automatically assess which image is more appealing than the others.

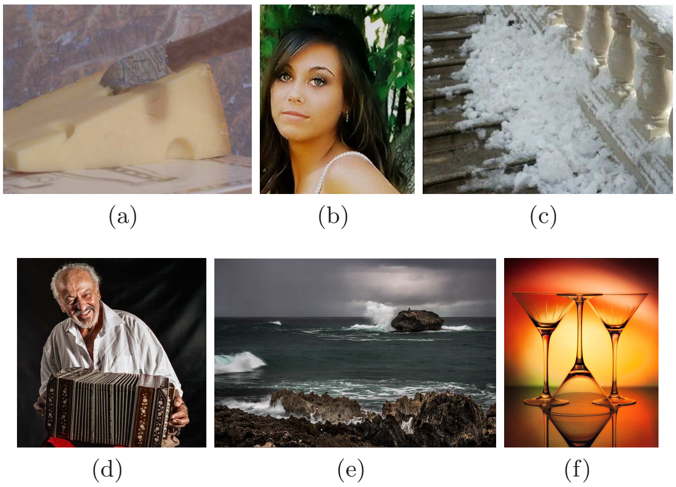


Fig. 1. Examples of low aesthetics quality photographs vs high aesthetics quality photographs. (a) Low aesthetics quality photo with blur. (b) Amateurish portrait with compression artifacts. (c) Example photo breaking the rule of thirds. (d) Portrait has professional lighting and color harmony. (e) Photo with well balanced composition. (f) Professional lighting and symmetry on a lovely colored picture.

Many early photo quality assessment approaches attempted to formulate those commonly accepted rules empirically, such as the methods described in [5–8]. The potential problem of this type of methods is that the models are designed based on authors' intuition and only global features are selected. In order to address the shortcoming of relying on global features only, some researchers focused their studies on local features which show better performance to predict aesthetics quality, as illustrated in [9–12]. Recently, researchers explored methods to encode photographic criterion indirectly into a learning phase, which discovered the relationship between aesthetics quality and underlying features automatically [13, 14].

Amongst the existing features used for photo aesthetics assessment, color harmony has been considered as one of the most important factors which has a significant influence on photo aesthetics but has been ignored for a long time. Only from the more recently, a few studies were applied on image aesthetics assessment by introducing some naive color harmony theories and models into the image quality classification/image harmonization tasks, which shows the effectiveness of the color harmony models [15–17]. However, those pioneer researches are still immature because the color harmony models employed are mostly based on the one which was found to have poor predictive performance [18], such as Moon-Spencer model used in [17], or the one which ignored partial important components for color harmony analysis, such as Matsuda model used in [15] where only Hue was used. More often, recent researches of color harmony theories in the image processing community have been focused on color selection and color design areas, where limited number of color combinations are suggested by the system according to some color harmony models [19].

Although some progresses of color harmony models have been made in the past years, it is still hard to apply them to the area of automatic image aesthetics assessment due to following reasons:

1. Most existing color harmony models are heuristically defined which can represent limited number of color harmony combinations. Although they are suitable for color design tasks, it is beyond its reach to assess the color compatibility of real world images with numerous color combinations.
2. Even plenty of color harmony models are proposed by researchers, they are defined in different color spaces and there is a lack of consensus between them. As these theories are various enough, nearly every color combination can be considered as harmonious if all models are considered [20].

In order to break through these limitations, researchers attempted to use machine learning approaches to "learn" compatible colors from large scale datasets. In [19], O'Donovan *et. al* trained a color harmony model from large datasets by using a least absolute shrinkage and selection operator (LASSO) regression approach, which can predict whether the user provided color combination is harmony or not. Inspired by this work, we applied a latent Dirichlet allocation(LDA) and Gaussian mixture model (GMM) based hierarchical learning approach to train a color harmony model dependent on large amount high quality images, and estimated the aesthetics quality of photos based on the trained color harmony model.

To the best of our knowledge, we are the first to reveal the process of how a color harmonious image is generated by learning this underlying principles from natural images. In summary, two main contributions are presented in this paper:

1. An generative, instead of empirical rule based, color harmony model is proposed by using large set of natural images to facilitate the task of image aesthetics assessment;
2. A hierarchical unsupervised learning model is proposed to learn the complex color combinations from photos. Based on this model, a principled probability based metric is also defined and applied for the aesthetics assessment task.

We organize the rest of the paper as follows. Section 2 is the review of related work, and section 3 describes the details of the proposed LDA and GMM based hierarchical color harmony model, including a brief introduction of these two models, and their implementation for our tasks. The experimental setup and analysis are provided in section 4. Section 5 concludes the paper.

2 Related Work

Color harmony, which is defined by Holtzschue as *"two or more colors are sensed together as a single, pleasing, collective impression"* [21], *"is one of the reputed daughters of Aphrodite, goddess of beauty, and this indicates that harmony is the province of aesthetics"*, stated by Westland in [22]. Due to the direct relationship between color harmony and image aesthetics, the exploration for the principles of color harmony has dominated the research of many artists and scientists.

Generally, most theories of color harmony follow the rule that multiple colors in neighboring areas produce a pleasing effect and thus can be roughly categorized into three types.

The first type of color harmony theories are originated from Newton's color theory, where hue circle is used to mathematically define different sets of color harmony principles. In [23], Itten suggested that a small number of colors uniform distributed on the hue wheel can be considered as harmonious. Derived from this idea, Matsuda designed a set of hue templates which defined ranges of harmony colors on the color circle [24].

The second category color harmony models are not only relied on hue information, but also introduce multiple features from color space to determine the color harmonious. In order to emphasize balance as a key factor of color harmony, Munsell [25] suggested that color harmony can be attained if colors with various saturation are in the same region of hue and value. To quantitatively represent Munsell's color harmony model, Moon and Spencer [26] proposed a model based on color difference, area and an esthetic measure, where color combination is harmonious when color difference is in the pattern of *identity*, *similarity*, or *contrast*.

The third category suggests that color harmony can be achieved when colors are similar in terms of hue or tone level. For example, in Natural Color System (NCS) which was detailed described by Hård and Sivik [27], colors are represented using the six elementary colors based on percepts of human vision, and the color combination can be classified according to *distinctness of border*, *interval kind* and *interval size*.

It should be noted that because the concepts of color harmony are highly dependent on nurture and culture, there are no obvious borderline between each type of classical color harmony theories. Many principles are shared by these theories and contradictory can also be found between them. Thus, in recent years, research interests were aroused in computer vision community to use machine learning approach to "learn" rules or patterns of color harmony from large data set based on images' statistical properties, such as method proposed in [19], where compatible color combinations were learnt from large amount rated images.

In this paper, we propose a hierarchical color harmony model to learn the underlying rules of compatible colors from high quality images, and evaluate this model by assessing the aesthetics quality of images. Our work is related to method presented in [19] in the sense of color harmony modeling and methods described in [17, 12] in the sense of image aesthetics assessment.

3 Hierarchical Color Harmony Model

"Pleasing joint effect of two or more colors" [28] is a prevalently accepted definition for the color harmony theories. Originated from this definition, a hierarchical color harmony model(HCHM) is proposed in this paper to learn those pleasing color combinations from images, which is used to predict the degree of harmony for unseen images. Initially, the HCHM learns the co-occurrence colors (color groups in our scenario) in images, followed by a GMM learning phase to encode the relations between color groups. By using this hierarchical structure, HCHM can model the complex color combination which represents the color harmony of the image.

3.1 Color Quantization

Prior to the color groups learning phase, each image is initially divided into small patches and colors are averaged within each patch. By quantizing each color patch using a color codebook, the mosaiced image can be represented by a set of "color words". Considering that human vision perception is more sensitive to the color with high perceived luminance, we use a non-linear quantization approach in HSV color space, which can be expressed as Eq. 1:

$$BIN = 1 + \sum_{i=2}^{L} i \times (i-1) \times q \tag{1}$$

where BIN is the total number of code words for hue-saturation-value (HSV) space, L means we partition the entire HSV space into L subspace evenly according to *value*, and then we divided each subspace using $(i-2) \times q$ radial lines and i circles. In this paper, we set $L = 10$ and $q = 4$.

Under this quantization scheme, the color space is coarsely partitioned in the region with low *value*, whilst the space is intensively separated in the region with high *value*, as illustrated in 2(a). In Fig. 2(b), an example image along with its corresponding mosaiced image and color words encoded image is shown.

Thus, a given image i can be represented as a set of color words $\mathbf{c} = \{c_1^{(i)}, \cdots, c_n^{(i)}\}$, where $n \in N$ and N is the total number of patches in image i.

3.2 LDA based Color Groups Learning

In the information retrieval and data mining area, Latent Dirichlet Allocation (LDA) is a widely used unsupervised approach to find patterns in unstructured

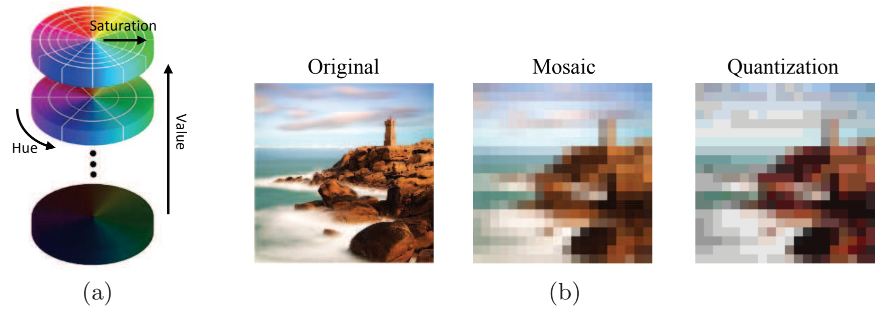


Fig. 2. An example of image mosaicing and quantization. (a) HSV cylinder is non-linearly divided into small regions. (b) A high quality image along with its mosaiced image and quantized image by using quantization codebook provided by Fig. 2(a).

data, which inherently has the capability to discover the semantic coherent item within large corpus. In the framework of LDA, each document can be represented by a distribution over topics, where each topic is a distribution over the vocabulary. In our scenario, by considering each quantized color as a word, LDA can be used to learn the coherent colors through topics, which correspond to color groups in HCHM, and to represent images' color property through the topics distribution.

By using LDA, the generative process to create an image can be formally described as below:

1. Given a set of K color topics:
   - Draw a multinomial color topic distribution over color codebook according to $\boldsymbol{\varphi}_k \sim Dir(\cdot \mid \boldsymbol{\beta})$ for each color topic $k \in \{1, \cdots, K\}$, where $Dir(\cdot)$ is a Dirichlet distribution, $\boldsymbol{\beta}$ is a V-dimensional Dirichlet parameter and V is the color codebook size.
   - A color topic matrix $\boldsymbol{\Phi} = \boldsymbol{\varphi}_{1:K}$ is formed, whose size is $V \times K$ and its element indicates the probability of the color word v given a topic k.
2. For each image i in the dataset:
   - Draw a parameter $\boldsymbol{\vartheta}^{(i)}$ that determines the distribution of the color topic. This is done by choosing $\boldsymbol{\vartheta}^{(i)} \sim Dir(\cdot \mid \boldsymbol{\alpha})$ for image i, where $\boldsymbol{\alpha}$ is a K-dimensional Dirichlet parameter. $\boldsymbol{\vartheta}^{(i)}$ is a K-dimensional parameter of a multinomial distribution, and $\boldsymbol{\vartheta}_k^{(i)}$ is the proportion of color topic k in image i.
   - To generate a color word (quantized image patch) $c_n^{(i)}$ in the image i:
     - Choose a color topic $z_n^{(i)} \in \{1, \cdots, K\} \sim Mult(\cdot \mid \boldsymbol{\vartheta}^{(i)})$, where $Mult(\cdot)$ is a multinomial distribution, $z_n^{(i)}$ is a color topic assignment and K is the total number of topics.
     - Choose a color $c_n^{(i)} \in \{1, \cdots, V\} \sim Mult(\cdot \mid \boldsymbol{\varphi}_{z_n^{(i)}})$, where V is the size of the color codebook.

With the quantized color words from the training set, the parameter $\boldsymbol{\vartheta}^{(i)}$ for each image and $\boldsymbol{\varphi}_k$ for each topic can be estimated by Gibbs LDA. In Fig. 3, we illustrate the color topic distribution of sample image in Fig. 2(b), which is inferred from our trained LDA. More details of LDA's training and inference can be found in [29].

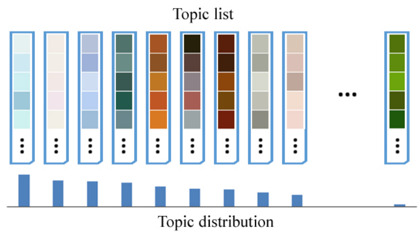


Fig. 3. Color topic distribution of sample image Fig. 2(b). Color topics are listed from left to right in the descending order according to their distributions within the sample image. In each topic, the top five color words are illustrated.

3.3 Applying GMM for Color Harmony Model

In order to model the dependency between topics for a given corpus, a Gaussian mixture model is learned on the top of LDA based on the color topic distribution $\boldsymbol{\vartheta}^{(i)}$. Theoretically, it can be shown that by using infinite number of mixtures, GMM can approximate every general continuous probability distribution to arbitrary precision.

Given a Gaussian mixture model, the probability of a topic distribution of image i in topic space can be described as:

$$p(\boldsymbol{\vartheta}^{(i)} \mid \xi) = \sum_{i=m}^{M} \omega_m \mathcal{N}(\boldsymbol{\vartheta}^{(i)} \mid \boldsymbol{\mu}_m, \boldsymbol{\Sigma}_m) \tag{2}$$

where $\mathcal{N}(\cdot)$ is a Gaussian probability density function:

$$\mathcal{N}(\boldsymbol{\vartheta}^{(i)} \mid \boldsymbol{\mu}_m, \boldsymbol{\Sigma}_m) = \frac{\exp\left(-\frac{1}{2}(\boldsymbol{\vartheta}^{(i)} - \boldsymbol{\mu}_m)^T \boldsymbol{\Sigma}_m^{-1} (\boldsymbol{\vartheta}^{(i)} - \boldsymbol{\mu}_m)\right)}{(2\pi)^{D/2} |\boldsymbol{\Sigma}_m|^{1/2}} \tag{3}$$

where $\boldsymbol{\vartheta}^{(i)}$ is the color topic distribution in topic space, which is obtained by applying LDA on the image i, parameters $\xi = \{\omega_m, \boldsymbol{\mu}_m, \boldsymbol{\Sigma}_m\}$ denote the weight, the mean and the covariance of the mth Gaussian distribution that satisfy $\sum_{m=1}^{M} \omega_m = 1$, and M is the number of mixtures. $\mathcal{N}(\boldsymbol{\vartheta}^{(i)} \mid \boldsymbol{\mu}_m, \boldsymbol{\Sigma}_m)$ reflects the probability of $\boldsymbol{\vartheta}^{(i)}$ being to the mth Gaussian distribution. For the correct estimation of Gaussian mixture model, the so-called Expectation-Maximization (EM) algorithm is used.

Normally, the likelihood $p(\boldsymbol{\vartheta}^{(i)} \mid \xi)$ can be used directly to measure the color harmony of unseen images. In our work, in order to fit the color harmony model into the image aesthetics assessment task, we use $\epsilon = \frac{p(\boldsymbol{\vartheta}^{(i)}|\xi^{+})}{p(\boldsymbol{\vartheta}^{(i)}|\xi^{-})}$ to represent the degree of color harmony of a given image, where GMM parameter ξ^{+} is trained by using high aesthetics quality images while ξ^{-} is trained with low aesthetics quality images. To an unseen high quality image, it should have higher probability with $p(\boldsymbol{\vartheta}^{(i)} \mid \xi^{+})$, whereas its probability on low quality images trained $p(\boldsymbol{\vartheta}^{(i)} \mid \xi^{-})$ should be lower. Then ϵ can represent the degree of color harmony.

4 Experimental Results

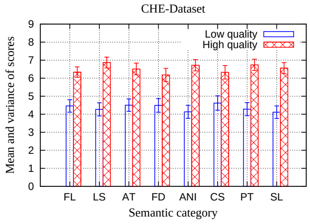
4.1 Dataset

In our experiments, we created a color harmony evaluation dataset (CHE-Dataset), which was a subset of AVA dataset [30], for our training and evaluation purpose. To the aim of aesthetics assessment, AVA dataset contains more than $250,000$ images which are categorized into over 60 groups. These images have plenty of meta-data, including a large number of aesthetics scores for each image, and semantic labels for groups. The quality scores of AVA dataset are based on various aesthetics aspects including color harmony, composition, subject, blur, etc. Only top ranked and bottom ranked images' scores are correlated with the degree of color harmony. So in order to meet the requirements of our color harmony model training and evaluation purpose, we followed the same method as [17] to select the top $2,000$ images and the bottom $2,000$ images based on their aesthetics scores for each category, respectively. All monochrome images were excluded from our dataset. To the high aesthetics and low aesthetics subsets of each category, images were evenly divided into a training set and a testing set. In this work, we collected the total number of $29,844$ images from eight different categories in AVA (a single image in our corpus may have multiple category labels), whose labels are: Floral, Landscape, Architecture, Food and Drink, Animals, Cityscape, Portraiture and Still Life.

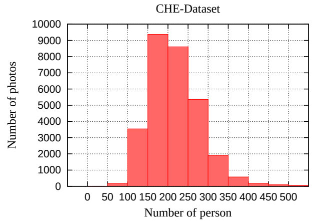
In Fig. 4, we illustrate the statistic properties of our dataset, where Fig. 4(a) shows the mean values and variances of high aesthetics quality subset and low aesthetics quality subset for each category. In 4(b), the number of individuals who provided scores for images in CHE-Dataset is shown. From this figure, we can see that most images in our dataset have plenty number of individuals (> 200).

4.2 System Analysis

To evaluate the performance of the proposed model, we firstly mosaiced each image and quantized image patches using the codebook introduced in Sec. 3.1. Then in the training phase, an LDA was trained by using 1000 high aesthetics quality images for each category. Further more, as described in Sec. 3.3, on the



(a)



(b)

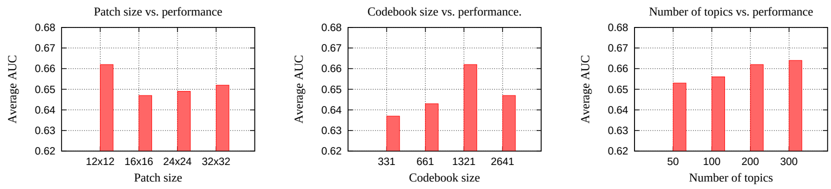
Fig. 4. Statistic properties of CHE-Dataset. (a) Box-plots of the means and variances of each category's scores. FL, LS, AT, FD, ANI, CS, PT and SL denote Floral, Landscape, Architecture, Food and Drink, Animals, Cityscape, Portraiture and Still Life, respectively. (b) The number of scores available for images in CHE-Dataset.

top of LDA we trained two GMMs with three Gaussian mixture components by using 1000 images with high aesthetics quality and 1000 images with low aesthetics quality, separately. In the testing phase, the color topic distribution $\boldsymbol{\vartheta}^{(i)}$ of test image i from a given category was computed based on corresponding LDA. Then the color harmony score ϵ was obtained based on the trained GMMs and the image with high ϵ score was classified as a high aesthetics image. In our experiments, 2000 test images from each category were used for the evaluation purpose.

In this experiment, different sets of parameters of patch size, quantization codebook size and topic numbers were investigated. As the target of the color harmony modeling is to assess the images aesthetics quality, we evaluated the performance of the proposed method using the average areas under the ROC curve (AUC) for the entire test set.

In Fig. 5(a), AUCs of four sets of patch sizes were evaluated and the best classification accuracy was achieved by using 12×12 patch size. From this figure, an interesting phenomenon can be observed that smaller or larger patch sizes can provide better performance than the patch size of 16×16. The reason is that the color harmony information can be encoded by LDA when smaller patches contain pure colors. Whilst, color compatibility information in image space can also be encoded in patch level with suitable patch size, which was revealed in Fig. 5(a).

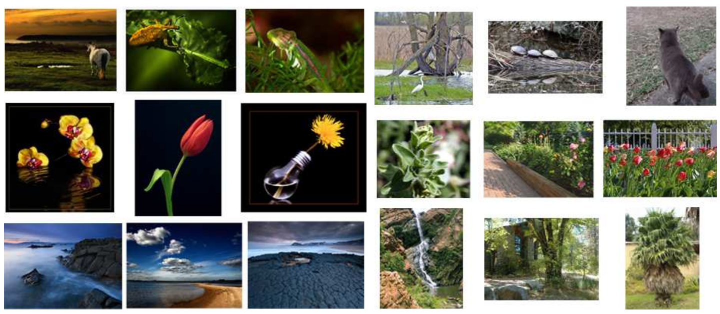
In Fig. 5(b) and Fig. 5(c), different values of the codebook size for quantization in HSV color space and the number of topics for LDA were examined, which showed better performance with codebook size of 1321 and topic number of 300 for aesthetics assessment. As the codebook size decreased, more colors were mapped to one code, which cannot effectively represent the color combinations before LDA and degrade the system performance. As the codebook size increased from some point, more and more similar color codes appeared in the same color topic, which cannot effectively represent the color combinations after LDA.



(a) (b) (c)

Fig. 5. AUC with different parameters. (a) AUC with various patch sizes which were used for image mosaicing. (b) AUC with different codebook size for quantization. (c) AUC with different color topic numbers for LDA.

To visually demonstrate the discriminate capability of the proposed color harmony model, we illustrated the top ranked images and bottom ranked images from three different categories: Animals, Floral and Architecture in Fig. 6. From this figure, we can see that although both top ranked and bottom ranked images contain rich colors, the proposed color harmony model can reveal the subtle difference between them and distinguish high quality images from low quality images.



(a) (b)

Fig. 6. High aesthetics quality images and low aesthetics quality images classified by the proposed system. (a) Top ranked images with high aesthetics scores. (b) Bottom ranked images with low aesthetics scores.

In order to discover what type of color combinations is learned by the proposed hierarchical color harmony model, we illustrated the top 10 color topics (groups) of each image category in Fig. 7, where 5 top ranked color words of each topic are shown. From this figure, we can see that the proposed model is capable to capture key color combination patterns for each category. For example, to the floral category, the flower related colors: purple, red, pink are among the top list through our system. And to the portraiture category, photographers tend to use dark colors as background, which are also revealed by the proposed system.

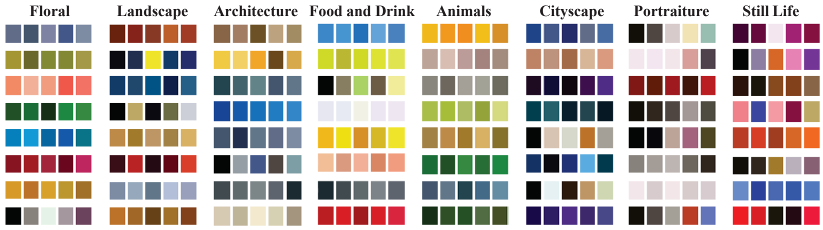


Fig. 7. Top-10 ranked color groups (topics) learned from LDA for each image category .

4.3 Comparison and Discussion

In our experiments, we compared the proposed unsupervised color harmony model to other three the state-of-the-art color harmony models for image aesthetics assessment task, which are briefly summarized below.

Moon and Spencer's color harmony model used a quantitative aesthetics measure $M = O/C$, where C represents complexity and O means order of a color combination. The quantitative definition of C and O can be found in [31].

Matsuda's model matched the hue distribution of an image against eight hue distribution templates, where the highest matching score was used to represent the degree of image's color harmony. In this model, the parameters of hue distributions were heuristically defined by aesthetics researchers.

Tang's color harmony model was based on Matsuda's hue template, where hue distributions of each template were learnt from a set of training images. Although the parameters of the Tang's hue template were adaptive for different type of images, it had the same drawback as Matsuda's method that only hue information was employed in the color harmony model.

In Fig. 8, the ROC curves of different approaches for each category were shown, where we can see that the proposed color harmony model provided more discriminant capability than other models to distinguish high aesthetics quality images from low aesthetics quality images.

To further analyze the performance gain of the proposed method, experimental results on different level of color complexities were illustrated in Fig. 9, where the images' aesthetics ranks of different methods were listed in corresponding tables, along with the user labeled ranks.

In Fig. 9(a), two sample images have a simple color combination (distribution) which can be modeled by hue templates (as shown by the corresponding hue wheels). To this type of images, both rule based color harmony models, such as Matsuda and M&S's models, and the learning based models, such as Tang's and the proposed methods, can effectively represent the harmony information contained in images.

To images shown in Fig. 9(b), whose color distributions cannot be accurately modeled by empirically defined hue templates, rule based methods predict relatively low aesthetics scores which causes those images had low aesthetics ranks.

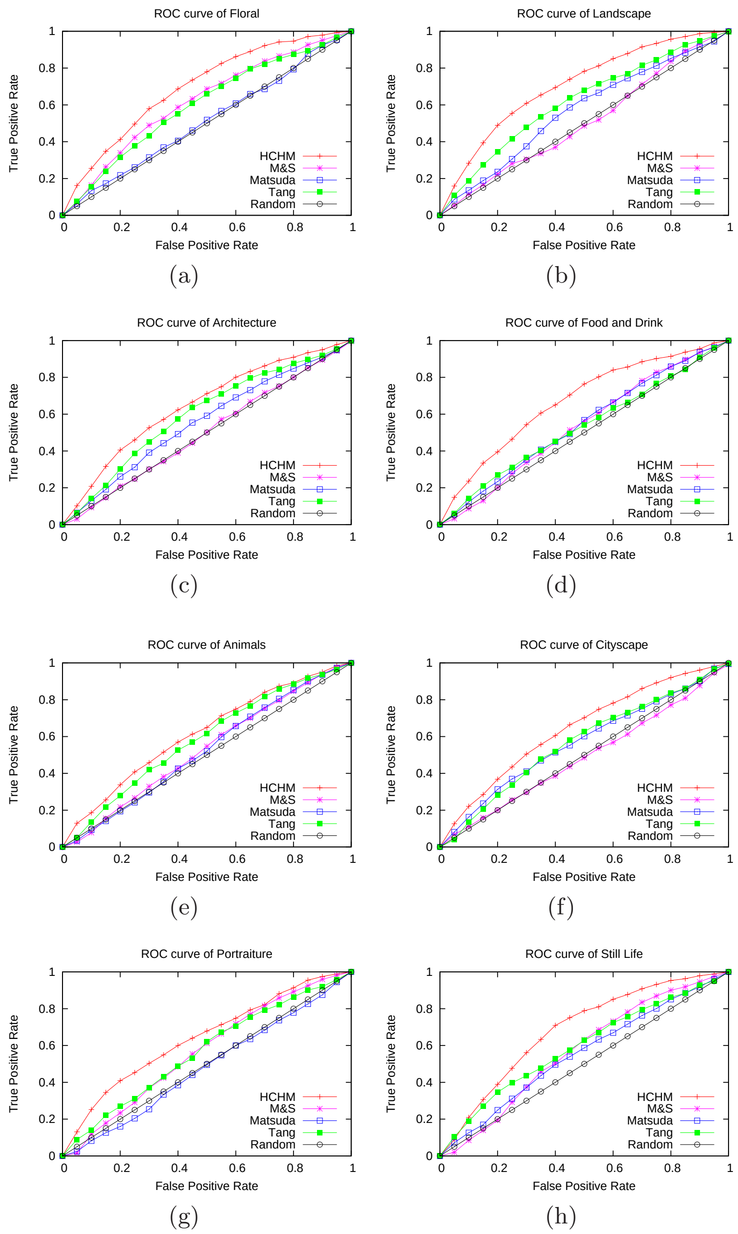
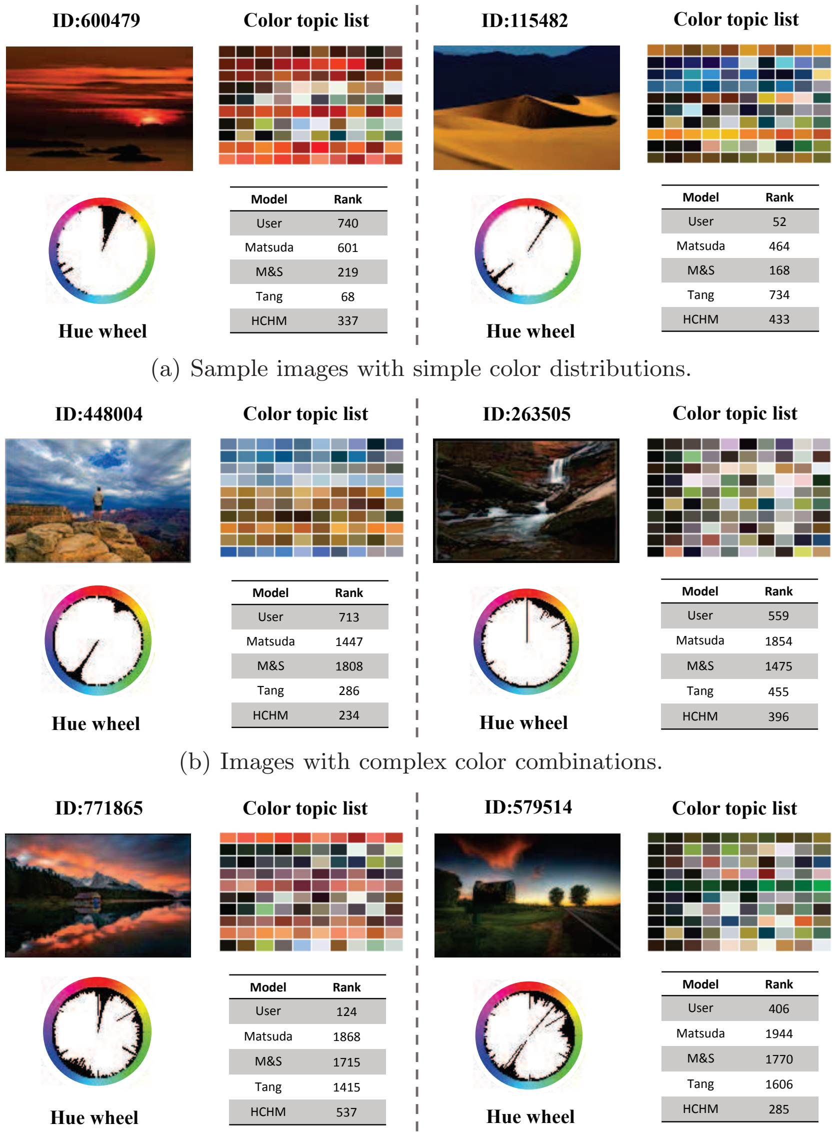


Fig. 8. ROC curves of image aesthetics assessment for each category.



ID:600479

Model	Rank
User	740
Matsuda	601
M&S	219
Tang	68
HCHM	337

ID:115482

Model	Rank
User	52
Matsuda	464
M&S	168
Tang	734
HCHM	433

(a) Sample images with simple color distributions.

ID:448004

Model	Rank
User	713
Matsuda	1447
M&S	1808
Tang	286
HCHM	234

ID:263505

Model	Rank
User	559
Matsuda	1854
M&S	1475
Tang	455
HCHM	396

(b) Images with complex color combinations.

ID:771865

Model	Rank
User	124
Matsuda	1868
M&S	1715
Tang	1415
HCHM	537

ID:579514

Model	Rank
User	406
Matsuda	1944
M&S	1770
Tang	1606
HCHM	285

(c) Images with wide range of color distributions that cannot be represented by heuristic based models.

Fig. 9. Comparison of image aesthetics assessment with different color complexities. To each sub-figure, the original color image along with its top 10 color topics (color groups) are illustrated in the top row. The corresponding hue distribution, the user labeled ranks and predicted ranks by different color harmony models are shown in bottom row of each sub-figure.

In the mean time, Tang's method can adjust the parameters of each hue template through the learning phase, which still provides reliable ranking score.

Given images with even more complex color combinations, as shown in Fig. 9(c), whose hue distributions cannot be modeled by any templates, even with adaptive parameters, the aesthetics scores provided by those methods which only consider particular color channel, even with learnt parameters from the training set, are much lower than truth. But by using the proposed color harmony model, color combinations are well encoded into the system which provides reliable aesthetics assessment ranking for test images.

Compared with rule based methods, the proposed color harmony model relies on the quality of the training dataset. From our experiments, we found that by using the data selection approach described in [17], the annotated scores of each photo still cannot precisely represent the color harmony degree for low quality images, where images with low quality scores may be harmonious but have other type of degradations, such as blur, bad composition, etc. This is the main reason for the most of misclassifications in our experiments. Thus, to build a dataset with more reliable scores measuring the color harmony degrees can overcome the problem of false alarm for the proposed color harmony model, which will be our future research focus.

5 Conclusions

In this paper, a hierarchical unsupervised learning approach is proposed to learn the compatible color combinations from large dataset. In this hierarchical framework, the LDA is adopted to learn the simple color groups, followed by the GMM training procedure to learn the combinations of color groups. By using this hierarchical structure, we can discover the harmonious colors, which facilitates the color selection for design industry. The HCHM can also encode the complex color combinations from the dataset images, which provides a feasible way to assess the aesthetics quality of natural images. Experimental results show the HCHM's capability of learning the complex color combinations.

The proposed method provides an example of learning the color harmony from natural photos, which is a promising way to increase our knowledge of color harmony.

Our future work includes to extend our method by taking the spatial relationship between colors in images and combining other features to improve the aesthetics assessment performance.

Acknowledgement. This work was supported by the National Natural Science Foundation of China (Grant No. 61273365 and No. 61100120) and the Fundamental Research Funds for the Central Universities (No. 2013RC0304).

References

1. Murray, N., Marchesotti, L., Perronnin, F.: Learning to rank images using semantic and aesthetic labels. In: 23th British machine and vision conference (BMVC).

(2012)
2. Geng, B., Yang, L., Xu, C., Hua, X.S., Li, S.: The role of attractiveness in web image search. In: Proceedings of the 19th ACM International Conference on Multimedia. (2011) 63–72
3. Wang, Y., Dai, Q., Feng, R., Jiang, Y.G.: Beauty is here: Evaluating aesthetics in videos using multimodal features and free training data. In: Proceedings of the 21st ACM International Conference on Multimedia. (2013) 369–372
4. Bhattacharya, S., Nojavanasghari, B., Chen, T., Liu, D., Chang, S.F., Shah, M.: Towards a comprehensive computational model for aesthetic assessment of videos. In: Proceedings of the 21st ACM International Conference on Multimedia. (2013) 361–364
5. Damera-Venkata, N., Kite, T., Geisler, W., Evans, B., Bovik, A.: Image quality assessment based on a degradation model. Image Processing, IEEE Transactions on **9** (2000) 636–650
6. Li, X.: Blind image quality assessment. In: Image Processing. 2002. Proceedings. 2002 International Conference on. Volume 1. (2002) 449–452
7. Datta, R., Joshi, D., Li, J., Wang, J.Z.: Studying aesthetics in photographic images using a computational approach. In: 9th European Conference on Computer Vision. Volume 3953 of Lecture Notes in Computer Science. (2006) 288–301
8. Ke, Y., Tang, X., Jing, F.: The design of high-level features for photo quality assessment. In: Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on. Volume 1. (2006) 419–426
9. Li, C., Chen, T.: Aesthetic visual quality assessment of paintings. Journal of Selected Topics in Signal Processing **3** (2009) 236–252
10. Luo, Y., Tang, X.: Photo and video quality evaluation: Focusing on the subject. In: Proceedings of the 10th European Conference on Computer Vision: Part III. (2008) 386–399
11. Luo, W., Wang, X., Tang, X.: Content-based photo quality assessment. In: Computer Vision (ICCV), 2011 IEEE International Conference on. (2011) 2206–2213
12. Tang, X., Luo, W., Wang, X.: Content-based photo quality assessment. Multimedia, IEEE Transactions on **15** (2013) 1930–1943
13. Marchesotti, L., Perronnin, F., Larlus, D., Csurka, G.: Assessing the aesthetic quality of photographs using generic image descriptors. In: Computer Vision (ICCV), 2011 IEEE International Conference on. (2011) 1784–1791
14. Marchesotti, L., Perronnin, F.: Learning beautiful (and ugly) attributes. In: 24th British machine and vision conference (BMVC). (2013)
15. Cohen-Or, D., Sorkine, O., Gal, R., Leyvand, T., Xu, Y.Q.: Color harmonization. ACM Trans. Graph. **25** (2006) 624–630
16. Tang, Z., Miao, Z., Wan, Y., Wang, Z.: Color harmonization for images. Journal of Electronic Imaging **20** (2011) 023001–023001–12
17. Nishiyama, M., Okabe, T., Sato, I., Sato, Y.: Aesthetic quality classification of photographs based on color harmony. In: Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on. (2011) 33–40
18. Pope, A.: Notes on the problem of color harmony and the geometry of color space. Journal of Optical Society of America **34** (1944) 759–765
19. O'Donovan, P., Agarwala, A., Hertzmann, A.: Color compatibility from large datasets. ACM Trans. Graph. **30** (2011) 63:1–63:12
20. Schloss, K., Palmer, S.: Aesthetic response to color combinations: preference, harmony, and similarity. Attention, Perception, & Psychophysics **73** (2011) 551–571
21. Holtzschue, L.: Understanding Color: An Introduction for Designers. Fourth edn. John Wiley & Sons, Inc. (2011)

22. Westland, S., Laycock, K., Cheung, V., Henry, P., Mahyar, F.: Colour harmony. Journal of the International Colour Association **1** (2007) 1–15
23. Itten, J.: The Art of Color: The Subjective Experience and Objective Rationale of Color. John Wiley & Sons (1997)
24. Matsuda, Y.: Color Design. Asakura Shoten (1995)
25. Munsell, A.H.: A Grammar of Color: A Basic Treatise on the Color System. Van Nostrand Reinhold Co (1969)
26. Moon, P., , Spencer, D.E.: Geometric formulation of classical color harmony. Journal of the Optical Society of America **34** (1944) 46–50
27. Hård, A., Sivik, L.: A theory of colors in combination descriptive model related to the ncs color-order system. Color Research & Application **26** (2001) 4–28
28. Holtzschue, L.: Understanding color: An introduction for designers. In: John Wiley & Sons, Inc. (2011)
29. Heinrich, G.: Parameter estimation for text analysis. Web: http://www. arbylon. net/publications/text-est. pdf (2005)
30. Murray, N., Marchesotti, L., Perronnin, F.: Ava: A large-scale database for aesthetic visual analysis. In: Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on. (2012) 2408–2415
31. O'Connor, Z.: Colour harmony revisited. Color Research & Application **35** (2010) 267–273