Neurocomputing 201 (2016) 82–91

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

Image color harmony modeling through neighbored co-occurrence colors

Peng Lu [a,*], Xujun Peng [b], Caixia Yuan [a], Ruifan Li [a], Xiaojie Wang [a]

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

ARTICLE INFO

Article history:
Received 23 May 2015
Received in revised form
11 August 2015
Accepted 15 March 2016
Communicated by Luming Zhang
Available online 21 April 2016

Keywords:
LDA
Color harmony
Esthetics
Linear regression

ABSTRACT

The traditional color harmony models for the photo esthetics assessment, such as Moon & Spencer's model and the adaptive hue template based approach, only utilize the distribution of (co-occurrence) colors based on heuristic rules or principled probability based metrics, where the spatial relationships between compatible colors are ignored. In this paper, we propose a discriminant learning approach to train a color harmony model based on Latent Dirichlet Allocation (LDA), which constrains the LDA's training by considering the spatial distances between harmonious colors. Our main contributions are: (1) employing spatial relationship between colors as the feature to build color harmony model, (2) proposing a consolidated framework to supervise the training phase of LDA, where the distribution and spatial information of colors and the target esthetics scores of images are used under this framework, and (3) designing an efficient algorithm to solve the alternating optimization problem for proposed color harmony model. The experimental results illustrate that compared to the existing color harmony models, the proposed method provides more reliable harmony scores to assess the photos' esthetics quality.

1. Introduction

In the past decade, we saw incredible demands in image retrieval for both online and off-line applications, such as web search engine and photo album manipulation. Besides the semantic relevance, users generally prefer more appealing images retrieved by image search engines or in photo database [1–4]. The importance of automatic esthetics assessment of images which is consistent with human perception is therefore widely recognized.

Although there have been many studies focusing on the image esthetics assessment, it is still a challenging problem due to two reasons. From the subjective point of view, different people may have various tastes of the same image because of different background of cultures, education, and personal moods, which define ambiguous esthetics. From the objective point of view, the beauty of a photograph depends on many factors, including composition, sharpness, proper lighting and contrast, as well as special photographic techniques.

To address various aspects of esthetics quality, the existing research mainly focused on extracting more advanced features and utilizing more sophisticated learning algorithms. In [5], Datta et al. suggested a computational approach where a set of color related and texture and composition based features were extracted and used in a support vector machine (SVM) to distinguish between high and low quality of esthetics. Based on the discovery that different esthetics criteria were taken by professional photographers with various types of photos, Tang et al. proposed a content-based photo quality assessment method which extracted objects from photos and applied different features for different objects prior to esthetics prediction [6]. To avoid the disadvantages of hand-crafted features, Lo et al. used color distributions of images in the HSV color space, along with other layout composition, edge composition and global texture features in SVM to accomplish esthetics assessment tasks [7]. Similarly, in [8], Wang et al. investigated a large set of low-level and mid-level features which were fed into SVM classifiers with combined χ^2 kernels and RBF kernels. Other researchers attempted to use multichannel local and global structural features to evaluate image's esthetics quality, where the spatial arrangements of image regions (spatially adjacent graphlets) were considered to create a probabilistic graphical model [9]. Again, in Zhang et al.'s esthetics assessment work, local esthetics descriptors were projected into a low-dimensional sematic space to describe the types of visual features [10,11]. In [12], Guo et al. utilized both hand-craft features and visual semantic features which represented the image topics to evaluate the esthetics quality of images. In recent studies, biological inspired features also attracted researchers' attention to assess the esthetic quality of medias [13].

* Corresponding author.

http://dx.doi.org/10.1016/j.neucom.2016.03.035

Although our perception of photo esthetics depends on many factors, and most of them are culture and context related, it is impossible to perceive and feel about an image without considering one of the most important features: harmonious colors. Color harmony, as stated in [14], “is one of the reputed daughters of Aphrodite, goddess of beauty, and this indicates that harmony is the province of esthetics”, which discloses the inherent relationship between image esthetics and color harmony. Thus, the exploration for the principles of color harmony is important to esthetics assessment of the images.

Recently, many researchers started to use existing color harmony models to evaluate images' esthetics quality through machine learning techniques. Inspired by the work in [15], where color harmony information was used to convert the original image into a more appealing color space, Nishiyama et al. used Moon–Spencer model in a bag-of-color-patterns framework to classify images based on their esthetics qualities [16]. In [17], Matsuda based color harmony templates were successfully applied for image quality assessment and editing. In [18], the same type of harmony templates were also used to enhance images’ visual harmony, where the harmonization was treated as an optimization problem. By utilizing color harmony model proposed by Ou [19], Gao et al. designed a computational model to evaluate the color quality of visible and infrared fusion images [20]. As revealed by mentioned works, applying existing color harmony models shows a promising prospect for esthetics quality assessment and enhancement tasks.

In this paper, by integrating the neighboring spatial relationship of compatible colors into a discriminant learning framework, we focus our work on discovering the underlying principles of generating harmonious colors. The rest of the paper is organized as follows: Section 2 provides a brief review of related work for color harmony theories. Section 3 proposes the framework for our color harmony model. Section 4 shows the mathematic details of the proposed spatial prior and linear regression constraint LDA (SPLRC-LDA) model. Section 6 illustrates the experimental setup and result. We conclude our work in Section 7.

2. Related work

The early research into the color harmony model can be traced back to the pioneering work on color theory, where different sets of color harmony principles were developed based on color-order systems, such as Matsuda's model which described the range of harmonious colors on the color circles [21]. Instead of using only the hue component in color circles, researchers also employed multiple features from color space to model the color harmony, such as Munsell color harmony model [22]. To quantitatively represent Munsell color harmony model, Moon and Spencer [23] proposed a method that utilized color difference, area and an esthetics measure to estimate the degree of color harmony.

Despite progress of color harmony models had been made, the traditional approaches mostly relied on heuristically defined rules or provided poor predictive performance for esthetics assessment [24]. To overcome the disadvantage inherited from traditional approaches, researchers tended to exploit machine learning approaches to “learn” color harmony models from large sets of images. One trend is to adapt existing color harmony models’ parameters based on training data, such as the method proposed by Tang [25]. In [26], Kuang et al. extended Tang's work to adaptively adjust the number of sectors for each hue template by mean shift based method. In [27], Chamaret and Urban used the arc-length distance to measure the similarity between a color image and a Matsuda harmonious hue template, which was then combined with other perceptual features by pooling approach to estimate image's quality. Another type of color harmony model research is to employ statistical learning method to obtain the color harmony model from training samples. The state-of-the-art works include the approach proposed in [28], where a color harmony model was trained through a set of high quality images based on LDA. In their extended work, Lu et al. demonstrated that by discovering the underlying color topics among images with different tags, the compatible colors can be found and used to predict the esthetics quality of images effectively [29]. By modeling classical color harmony theories as the prior knowledge, Lu et al. also proposed a Bayesian framework to integrate Matsuda/Munsell color harmony models into LDA's training to achieve better performance for esthetics assessment [30].

Although promising results had been provided by classical LDA-based color harmony models, they had two potential problems by their nature: (a) The existing LDA-based models only consider the distribution of compatible colors but ignore their spatial relationship in images. (b) LDA is a generative modeling approach which is lack of discriminability to fulfil the esthetics assessment task if we consider it as a classification problem.

As revealed in [31], “adjacent color regions can have strong effects on their neighbor's perceived color”, which shows that the pleasing effect of photos can only be obtained with multiple compatible colors in neighboring areas. Although few research had been done to utilize colors’ spatial information in LDA to assess images’ esthetics, many researchers used spatial configuration of regions to model the scene images by adopting topic model based approaches. For example, in [32], Lin and Xiao used a composite of regions to represent images, where each region was associated with a semantic topic. By focusing on organizing the local patches of image into multiple discriminative subgraphs, Hu and Nie proposed a framework to categorize aerial images relied on subgraphs’ spatial relationships [33]. The advantage of this method is that the dependencies between both locations and topics are taken into consideration in the proposed generative model. In order to overcome the drawback of the traditional probabilistic Latent Semantic Analysis (pLSA) and LDA, which assume each region of the image is independently generated, Cao and Li proposed a spatially coherent latent topic model (Spatial-LTM) to achieve better classification results [34]. For this approach, each image was modeled in a hierarchical manner by over-segmenting image regions of homogeneous appearances and all patches within the same region shared a single latent topic, which constrained the model with the spatial coherency. By using the similar idea, Niu et al. designed an image classification system which modeled the statistics of the appearances and locations of different regions for image objects, where a supervised topical model was employed to encode the spatial properties into the system [35].

Despite the mentioned LDA-based color harmony models’ capability to discover coherent colors from images, the discriminative information is not taken into account during those color harmony models training phase. Thus, to introduce the discriminability into LDA, researchers proposed supervised topic models by introducing the label information of each document/image during the training phase, which provide more reliable classification performances. In [36], a semi-supervised topic model was described, where both semi-supervised properties and supervised topic model were integrated into a regularization framework simultaneously. By considering both observed label frequencies and label dependency, Li et al. designed a supervised topic model to solve the multi-label classification problems [37]. In [38], Zhu et al. proposed a framework which combined the max-margin prediction criterion with the hierarchical Bayesian topic models. This framework was claimed to be more discriminative and suitable for classification and regression tasks.