Neurocomputing 172 (2016) 244–252

Contents lists available at ScienceDirect

Neurocomputing

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

Synthesized computational aesthetic evaluation of photos

Weining Wang [a,1], Dong Cai [a,1], Li Wang [a], Qinghua Huang [a,*], Xiangmin Xu [a], Xuelong Li [b]

[a] School of Electronic and Information Engineering, South China University of Technology, Guangzhou 510640, Guangdong, China
[b] The Center for Optical Imagery Analysis and Learning (OPTIMAL), State Key Laboratory of Transient Optics and Photonics, Xi'an Institute of Optics and Precision Mechanics, Chinese Academy of Sciences, Xi'an 710119, Shaanxi, China

A R T I C L E I N F O

Article history:
Received 26 November 2013
Received in revised form
11 November 2014
Accepted 24 December 2014
Available online 12 May 2015

Keywords:
Aesthetic evaluation
Feature extraction
Classification
Mobile application
Photo

A B S T R A C T

Assessing aesthetic appeal of images is a highly subjective task which has attracted a lot of interests recently. It is an interdisciplinary subject related to art, psychology, and computer vision. In this paper, we systematically study prior researches of feature extraction in this area, and category them into four groups, low level, rule based, information theory, and visual attention. In each group, the effectiveness and limitations of existing features are examined. Based on the analysis, we propose a comprehensive feature set, which include 16 novel features and 70 well proved features. With this feature set, we build the system under machine learning scheme consisting of an SVM based classifier to estimate if an image is *high aesthetic* or *low aesthetic*. The experiments are conducted on public datasets show that our comprehensive feature set outperforms conventional models that concentrate mainly on certain types of features. The combination of our features produces a promising classification accuracy of 82.4% and a good performance comparable to aesthetic rating of human. Finally, we implemented the proposed evaluation system on mobile devices. It can provide real-time feedback to help users capture appealing photos.

1. Introduction

Nowadays, multibillion consumer photographs have been produced with the popularity of social networks and smartphones. The demand for large scale of photo management is imminent [42]. A few of studies addressed this problem by resorting to statistical approaches which learn from metadata of images [2,3] or analyzing user's online photo albums at social network for photo storytelling [4].

Aesthetics has been discussed by philosophers for ages [1]. In the world of photography, aesthetics plays a major role by helping people appreciate and judge the beauty of photos. With respect to judging image beauty of a large collection, computational aesthetic evaluation is valuable and in great demand to manage the images not only for amateurish users but also professionals, and thus free them from the tedious pictures management work. For example, it can help people to select exciting and beautiful images from large collections and filter out the unappealing one at the same time. It can also be exploited in picture editing for enhancement by providing feedbacks based on certain aesthetic rules [5]. Since the functions of image aesthetics are universally applied to many fields including design [6], advertisement, photography [14,22] etc, they show a promising prospect in human's daily life.

Computational aesthetic evaluation of images appears as an interdisciplinary subject that is pertinent to art, psychology, computer science and so forth. It has recently attracted a lot of interest [5]. Computational aesthetic evaluation aims to simulate human visual system and perception to perform aesthetic judgment on images automatically. However, due to its subjectivity and complexity, it remains a highly challenging task. Despite of the difficulties, many researchers have attempted to solve these problems. The framework of computational visual aesthetics usually comprises two stages [7–9,12]. First, image features are modeled based on certain theories or assumptions related to human aesthetic appreciation. Second, the aesthetic predictive model of classification is built by machine learning from image features and subjective scores of humans. Finally, the aesthetic results are automatically given through the models. The effectiveness of the models is verified by comparing the predicted results with judgments or ratings of humans.

In early work [7–9], some common criteria on photography were used as guidelines to extract global visual features, such as color distribution, average brightness, wavelet texture. Datta et al. [9] initially realized the image classification of *low aesthetic* versus *high aesthetic*, and demonstrated the feasibility of computational aesthetics analysis. Ke et al. [7,8] proposed a set of principled methods in terms of photography and underlined the significance of simplicity

* Corresponding author. Tel.: +86 15013262380.
E-mail address: qhhuang@scut.edu.cn (Q. Huang).
[1] Weining Wang and Dong Cai have made the same contributions to this paper.

http://dx.doi.org/10.1016/j.neucom.2014.12.106
0925-2312/

and clarity for a professional photo. Their model could distinguish professional photos from snapshots, and achieved promising results. Obrador et al. [13] explored the role that image composition plays in image aesthetic evaluation. They proposed a set of composition templates that approximate photography guidelines.

Some researchers studied the visual art images [11,12]. Li et al. [11] revealed that a beautiful painting shares many aspects in common and can be computationally measured. They explored some characteristics related to artistic knowledge, such as color templates, blur effect. Li et al. [12] proposed a fuzzy approach to extracting aesthetic semantic feature of paintings as well.

Recently, efforts have been made to model the aesthetic measurements from information-theoretic perspective. Birkhoff wrote the first quantitative theory of aesthetics [15]. He claimed that images which are easy to process but visually complex to have higher aesthetic value. Until recent years, Rigau et al. [19] used different compression approaches to measure complexity of images based on information theory, and explored their relations with image aesthetics. Then Romero et al. [17] tried to use these complexity features on image aesthetics evaluation and got impressive results.

Other studies shed light on visual psychology. They looked into the connection between human's attention and image aesthetics. Wong et al. [18] stressed the importance of salient regions because the key role of visual attention in the evaluation of image aesthetics had been verified by several psychological studies. They extracted the salient region by using a computational visual attention model, which result in an enhanced performance of their model.

In recent years, a few of researchers have attempted to put their work into practical use. Datta et al. [20] set up the first automatic aesthetics rating website ACQUINE for photographic pictures. Recently, Lo et al. [23] designed an application on portable device which can display some basic and intuitive aesthetic scores for corresponding visual properties such as saturation, color, and contrast. Although above studies presented applicable system for users, some of them relied on network transitivity [20] or only provided limited aesthetic measurements [23].

Although researchers made significant progress in this area, aesthetic evaluation is still just started and has great space to improve. Various types of features are proposed, but not well synergized yet. Only part of image features and aesthetic findings are used by aesthetic evaluation. These inspire us to design new features and well integrate existing features for better evaluation.

The rest of the paper is organized as follows. Section 1.1 outlines the main structure of our aesthetics evaluation study. In Section 2, we introduce the features extraction for our evaluation work. Section 3 elucidates the machine learning methods we use for establishing our models. Section 4 presents the datasets and methodology we adopt and evaluate the performance of our system. We also compare our method with prior work and analyze the importance of different features in our model. Section 5 shows the application of our system used in mobile device. Finally, we conclude our work and highlight future research in Section 6.

1.1. Overview of our work

Our work has four parts, the first part is the feature set design, which is the most important part of our work. We analyze and test the existing well used features, and get an effective group of features to build our model by keeping some classic features, extending and revising some newly developed features, and designing some novel features. We aim to cover valuable features from different aspects so as to make the model generalizable across a wide variety of images. Thus we realize the classification model of aesthetic evaluation with improved performance by using these features.

After feature extraction, we combine the features and establish the aesthetic evaluation model by using machine learning. We use a Support Vector Machine (SVM) classifier to classify images into *high aesthetic* group and *low aesthetic* group according to their aesthetic appeal.

We also compare our method with prior work and analyze the importance of different features in our model. Moreover, our method has been put into practical application. We manage to establish our system on mobile device as a useful tool for users to take more appealing photos.

2. Feature extraction

In the field of computer vision, image features play a key role in aesthetic evaluation. Diverse features are modeled through computer vision techniques from different perspectives. Some features have been used and proved to be effective by previous studies. Some features are borrowed from other computer vision applications. In addition, we introduce some newly adopted features to enhance the performance of our model.

In total, we have extracted 86 features (including 16 newly adopted features) from an image to describe its aesthetic appeal as listed in Table 1. Among the 16 new features, some features are newly designed, the other features have been improved in calculation algorithms for better results. This table shows the names of features as well as the categories they belong to. They are designed based on not only intuitive basic computer vision guidelines, but also the rules of photography, art, information theory, and assumptions on visual psychology.

(1) Low-level features include the fundamental image elements such as colors, textures. Existing studies had verified the effectiveness of these features.
(2) Rule based features, are designed according to some rules of thumbs in certain domains of photography and art. People have accumulated a lot of experiences in photography and art for a very long time, and use these rules consciously when taking photos. Researchers design methods to extract features using these experiences. They capture color harmony, lighting, clarity and composition related to photographic aesthetics.
(3) Information theory based features, measure the complexity and order of images. The set of features conceptualized Birkhoff's theory measure objects from an informational viewpoint. Since the information theory and Birkhoff's theory can model and measure how human perceive images, these features are found to be useful in aesthetic evaluation recently.

Table 1
Proposed features in our method.

Feature	Meaning of feature	Category	Description
$f_1 \sim f_9$	Color moment	Low-level	
$f_{10} \sim f_{41}$	Gabor wavelet texture	Low-level	
$f_{42} \sim f_{53}$	Daubechies wavelet texture	Low-level	
$f_{54} \sim f_{56}$	Tamura texture	Low-level	
$f_{57} \sim f_{60}$	Color template[a]	Rule-based	Improved
$f_{61} \sim f_{64}$	Dark channel[a]	Rule-based	New design
$f_{65} \sim f_{67}$	Depth of field	Rule-based	
$f_{68} \sim f_{70}$	Image complexity of color[a]	Information theory	Improved
f_{71}	Image complexity of texture[a]	Information theory	New design
f_{72} , f_{73}	Processing complexity	Information theory	
$f_{74} \sim f_{82}$	Regional color moment[a]	Visual attention	Improved
$f_{83} \sim f_{86}$	Regional Gabor texture[a]	Visual attention	Improved

[a] Represents newly adopted features.