10-2014

A Multi-dimensional Image Quality Prediction Model for User-generated Images in Social Networks

You YANG
Huazhong University of Science and Technology

Xu WANG
City University of Hong Kong

Tao GUAN
Huazhong University of Science and Technology

Jialie SHEN
Singapore Management University, jlshen@smu.edu.sg

Li YU
Huazhong University of Science and Technology

DOI: https://doi.org/10.1016/j.ins.2014.03.016

Published in Information Sciences, 2014 Oct, 281, pp. 601-610.
http://doi.org/10.1016/j.ins.2014.03.016

A multi-dimensional image quality prediction model for user-generated images in social networks

You Yang [a,b], Xu Wang [c], Tao Guan [d,*], Jialie Shen [e], Li Yu [a]

[a] Department of Electronics & Information Engineering, Huazhong University of Science & Technology, Wuhan 430074, China
[b] Department of Automation, Tsinghua University, Beijing 100084, China
[c] Department of Computer Science, City University of Hong Kong, Hong Kong
[d] School of Computer Science & Technology, Huazhong University of Science & Technology, Wuhan 430074, China
[e] School of Information Systems, Singapore Management University, Singapore

Keywords:
Image quality prediction
User generated content
Multi-dimensional signal processing
Social networks

ABSTRACT

User-generated images (UGIs) are currently proliferating within social networks. These images contain multi-dimensional data, including the image itself, text and the social links of the owner. UGIs can be utilized for self-presentation, news dissemination and other purposes, and the quality of the image should be able to reveal these social functionalities. However, it is challenging to predict UGI quality utilizing existing models, such as image quality assessment, recommender systems or others, because these models have difficulties processing multi-dimensional data simultaneously. To address this problem, we propose a multi-dimensional image quality prediction model for UGIs in social networks. In this model, we build two sub-models for presentation measurement and distortion measurement. The text (i.e., tags and comments), social links and UGIs are processed by these two models separately, and the results of the models are pooled to obtain a final quality score. Both subjective and objective experiments are then arranged for ground truth data and performance assessment, respectively. Participants are asked to make judgments about 55 UGIs randomly selected from social networks, and the ground truth dataset is based on these subjective experiments. The objective experiments are performed to verify the performance of our model. The results indicate that the Pearson correlation parameter between the predicted score and the ground truth data is 0.5779, which suggests that the proposed model can be implemented to predict image quality in practical environments.

1. Introduction

User-generated content in social networks, especially the images or photos uploaded by end-users (i.e., user-generated images (UGI)), provides new opportunities for web publishing and media production [8,10,18,32,56]. These UGIs are re-editable, accessible and affordable for ordinary people in social interactions on the web. Actually, UGIs are special multimedia that provide multi-dimensional data, including the image itself as well as the comments, tags and social circles of the owner. These multi-dimensional data provide abundant information to viewers about the UGI and the image provider. Due to their convenience, billions of UGIs are uploaded and published on the web. For example, Facebook currently contains

* Corresponding author.
E-mail address: qd_gt@126.com (T. Guan).

100 billion image uploads from end-users, while Flickr stores 60 billion images and experiences a 20% annual increase [8,46]. These images can be utilized for self-presentation, news dissemination and other purposes, and the image quality should reveal the quality of its social functionalities. However, the prediction of UGI quality utilizing existing models, such as image quality assessment, recommender systems or others, is challenging because these models face difficulties in processing multi-dimensional data simultaneously. Moreover, social functionality is not included in the scope of these models [20,45]. Previously, only clinical methods were available for image preference measurements under controlled test environments, and these methods utilized carefully selected images rather than UGIs [23,30,63]. Therefore, the proliferation of UGIs requires a new approach to image quality prediction, which has considerable potential applications in image retrieval, recommendation systems and other fields [16,17,25,26,50,51,55].

Image quality prediction of a UGI is challenging because the uploaded image itself plays a multi-purpose role in social activities. Traditional images, those captured by cameras, record and reveal the real world as precisely as possible without social functionality. However, the main purpose of a UGI is social functionality. UGIs exist because the creators wish to present themselves to others in a virtual world [52]. Usually, such presentations are conducted through self-presence and self-disclosure [11,12,53]. Therefore, although the images captured by a camera record the real world, images re-edited by the provider reveal an aspect of his mind before publishing. Furthermore, UGIs are usually accompanied by text (i.e., descriptions, tags or title of the UGI) written by the provider for to better reveal his opinions and mind. To this end, the quality of the UGI should reveal the quality of self-presence to the user's social circle or the general public.

The special purpose of UGIs creates two challenges for image quality prediction of this type of image.

- Evaluation of personalized presentation. As mentioned above, the main purpose of a UGI is self-presentation in the virtual world. It is a challenge for researchers to find a relationship between quantified results generated by quality prediction of UGIs to the degree of self-presentation of the user.
- Measurement of unknown distortion. In addition to self-presentation, the intrinsic image quality also plays an important role in UGI quality. Traditional image quality assessment focuses on the measurement of distortions because the images are treated as visual signals. In these measurements, distortions are usually known, such as Gaussian noises and blurring effects. However, the UGI is edited by the provider. Therefore, unknown distortions may also be present with traditional distortions; therefore, traditional distortion measurement methods are unsuitable.

Existing methods may relate to the quality prediction of a UGI, such as data quality evaluation, item recommender systems and multiview machine learning. In data quality evaluation, the data are collected by users intentionally and consist of propositions that reflect reality. The collected data are necessary and useful for people's daily work, but the rapidly expanding volume of data can confuse consumers and hamper work efficiency. In this case, data quality is evaluated by its fitness for use by data consumers [6,36,42,47,59]. The purpose of data quality research is to optimize work flow and improve work efficiency. From this perspective, the purpose of data quality assessment is similar to UGI quality prediction. When UGIs are treated as data, the images can be evaluated utilizing the methods of data quality evaluation. However, traditional data quality evaluation focuses on economic data with inter-connections rather than images. Traditional evaluation is also problematic when using item recommender systems to assess the quality of UGIs. An item recommender system seeks to predict the rating or preference of a user for items (such as music, books, or movies) or social elements (e.g., people or groups) they had not yet considered employing a model incorporating the characteristics of an item (content-based approaches) or the user's social environment [2,7,24,48,49]. Actually, the recommender system is a retrieval system utilizing a user's profile. That profile, including favorite goods, selection habits, location and other useful information, can be compiled utilizing the historical purchase data of the registered user. Methods for modeling and analyzing this historical data are preference-based filtering methods, which include content-based recommendation, collaborative recommendation, and hybrid approaches [2,28,29,31]. Although UGIs can also be treated as social items, obtaining the potential viewer's profile, which is the core input of the recommender system, is difficult. Semi-supervised multiview machine learning methods have also been proposed to solve recent complex problems such as labeled and unlabeled joint feature processing [5,15,39–41,43,62]. Intuitively, the features of one domain are divided into different but compatible and uncorrelated views in multiview learning problems. The data are compatible if all examples are labeled identically by the target concepts of each view. At the same time, the descriptions of any given label are independent between two views. In these methods, non-negative matrix factorization is widely used for dimension reduction, and the learned results are similar in each view [14,19,21,22,33,35,38,57,58,61]. In our work, UGIs contains different types of data, including the image content, user comments, and even social links. All of these data can reveal UGI quality and can be treated as uncorrelated views. However, we cannot describe a UGI as high quality if it has little signal distortion but many negative comments. In other words, the classification of UGIs is a pooling process of the independent views of the data rather than a single view. Therefore, new methods should be developed to evaluate the quality of UGIs.

We propose a multi-dimensional image preference prediction model to evaluate the quality of UGIs on social networks. We claim that UGI quality is measurable, and the quality measurement is composed of presentation and distortion measurements. In this scheme, we build two sub-models for presentation measurement and distortion measurement. The text (i.e., tags and comments), social links and the UGI are processed by these two sub-models separately. Presentation measurement is used to evaluate the degree of self-presentation, which is the main purpose of UGIs in social networks. Distortion measurement is used to predict the perceptual quality of these images. Finally, these outcomes are pooled to obtain a final quality score. Therefore, our method processes the images, texts and social links as multi-dimensional data to obtain quality evaluations. We build a UGI quality database to provide benchmark data for model performance assessment. The UGIs are randomly selected from social networks, and participants are asked to provide ratings for these images. Then, objective scores of predicted quality are compared to the benchmarks, and the proposed model can be evaluated utilizing the database.

The rest of this paper is organized as follows. Section 2 provides a detailed discussion of our model. The subjective and objective experiments are discussed in Section 3. Section 4 concludes.

2. The proposed multi-dimensional quality assessment model

The proposed model of quality prediction considers all of the obtainable UGI data together to address the aforementioned challenges. The data of UGIs are a multi-dimensional, including the image, text (i.e., title, description and comments), and social circles of the image provider. Different types of data can be processed separately. For example, the image is characterized by pixel intensity, tags and comments by syntax and semantics, and social circles are described by nodes and links. The comments of the UGI and the social links of the provider are useful to understanding the interaction and feedback from the provider's social circles and can be used to evaluate the degree of presentation in the UGI. The pixel intensity of a UGI is usually compressed by JPEG encoder and the signal distortion is measurable in a no-reference manner. Therefore, the multi-dimensional data can be processed separately, and we describe the framework of the proposed model presented in Fig. 1.

As shown in Fig. 1, there are two sub-models within the framework. In the first sub-model, the multi-dimensional data are input into a multi-dimensional processing unit, where the texts and social links are used for presentation measurement. The pixel intensity of the UGI is simultaneously recorded for distortion measurement. The different types of data are used for the corresponding measurement methods yield a normalized score within the range [0, 10]. These normalized scores are then pooled to obtain a final quality score. Therefore, our model can be described by the following equation:

$$\varsigma(x,\mu) = f(\epsilon(x), \psi(\mu)) \tag{1}$$

where x is the number of comments, μ is the UGI, $\epsilon(\cdot)$ is the presentation measurement, $\psi(\cdot)$ is the distortion measurement, and $f(\cdot)$ is the pooling function. The details of each function will be discussed in the following subsections.

2.1. Presentation measurement

In our model, presentation measurement is based on text sentiment analysis from viewer comments and probability analysis of social links. As mentioned above, UGIs aim for self-presentation and reveal a user's mind or ideas. Sometimes, descriptions (i.e., tags) of the UGI are provided with the image. However, the comments from the UGI viewers are more valuable for evaluating the quality of UGIs because the comments reveal the viewers' attitudes. The quality of a UGI will be lower if fewer positive comments are given, regardless of how beautifully the image is edited or the description is written. These comments come from viewers that linked to the UGI provider. Therefore, the image quality will be higher if it receives more comments. However, whether a viewer prefers to post his comments is a stochastic event. We first use a random graph model to analyze the stochastic event. Then, we analyze the words in the comments to determine whether they have a positive sentiment.

2.1.1. Analysis of social links

We describe a graph model $G = (V, E)$ with vertex set V and edge set E, where $|V| = N + 1$ and $|E| = M$ denote the number of vertices and edges, respectively, to analyze the social links. In this model, the users are denoted by the vertices and the relationships between users are described by the edges. Specifically, we define a social link in a network as the single direction edge $\overrightarrow{e}$ from one user to another. In other words, we join two vertices (v_i, v_j) by $\overrightarrow{e}$ if v_j is a fan of v_i. These two users are called friends if they are both linked as a fan to their counterpart. Fig. 2 illustrates the conversion of the social links in

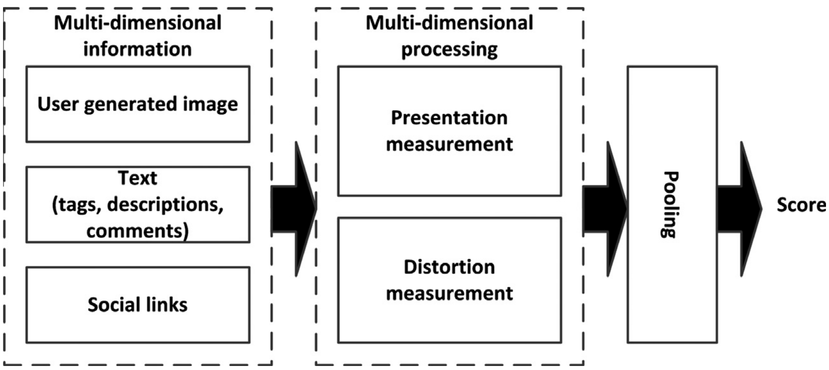

Fig. 1. The framework of the multi-dimensional image quality prediction model for UGIs in social networks.

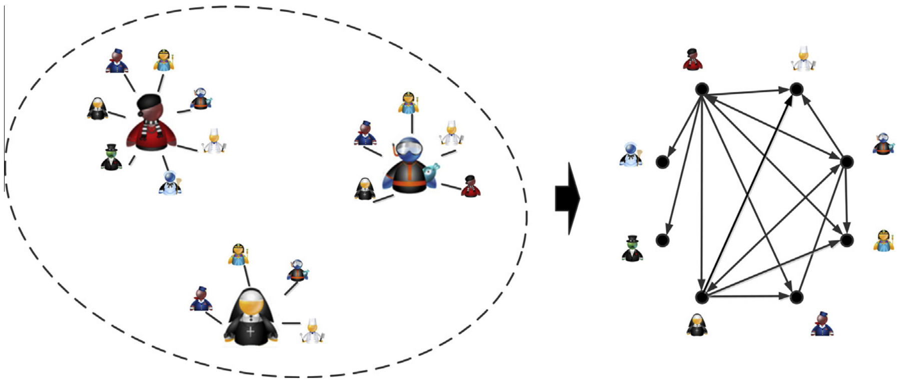

Fig. 2. Social links in networks and the corresponding graph model describing the relationships.

networks into a graph model. In this figure, the three selected users each have a different number of fans. For example, the man in red has seven fans, the man in black has five, and the lady in black has four. We use a vertex to denote a person and use a directional edge to denote the relationship of fan to obtain Fig. 2. The vertex that represents the man in red yields seven directional edges to all of other vertices, which describes his social links in the whole social network. Other vertices and edges provide similar meanings. If only one UGI provider v_i is selected, the graph model G_i for v_i is a single-directional graph, where all edges $\forall e \in E$ are with tail on v_i.

Based on the graph model G, we assign probabilities to edges to describe the stochastic event of comment posting. When a user v_i posts an image (i.e., a UGI) on his webpage, all his fans $V - \{v_i\}$ can visit and comment on this UGI. Suppose that any viewer $v_j(v_j \in V - \{v_i\})$ posting his comment to a UGI is not dependent on other viewers. In other words, it is a mutually independent event for different fans to comment on a UGI. Furthermore, we suppose that one viewer provides one comment, and duplicated comments are treated as a single comment. Therefore, we can utilize a bi-nominal distribution function to model our assumptions.

Let p be the probability that one fan provides one comments to one UGI; the probability that this user receives k comments from k fans is

$$P_k = \binom{N-1}{k} p^k (1-p)^{N-1-k} \tag{2}$$

where N is the number of vertices in $V - \{v_i\}$.

Let $x \in I^+$ be a positive integer and $x \leqslant N$ be possibility that this user receives at most x comments:

$$E(x) = \sum_{k=0}^{x} P_k \bigg|_{x \leqslant N-1} = (N-1-x)\binom{N-1}{x} \int_0^{1-p} y^{N-x-2}(1-y)^x dy \tag{3}$$

The function $E(x)$ in Eq. (3) is a monotonically decreasing function for both variables x and p with lower bound when $x \to \infty$. The monotonic behavior of $E(x)$ indicates the low probability that one UGI receives more comments when the number of users in the social network is fixed, regardless of how the fans favorite this image (i.e., described by p). However, a UGI must be of high quality if x is larger because more fans post their comments to this image, which is very difficult to achieve.

2.1.2. Analysis on viewer comments

The comments from viewer usually express the attitudes or moods of the image viewer, which may include love, like, happy, noncommittal, dislike, disgust, etc. These sentiments can be used to describe the quality of UGIs.

The question in comments based presentation measurement is how to convert the words in a comment to a value describing the quality of an image. There is a considerable gap between natural words and scale values. Fortunately, these comments can be processed utilizing sentiment analysis or opinion mining methods. Sentiment analysis or opinion mining refers to the application of natural language processing, computational linguistics, and text analytics to identifying and extracting subjective information from source materials [37]. Utilizing word sentiment analysis, the comment based quality of the UGI is measurable.

Comment-based quality measurement required a determined number of mood types. Traditionally, there are only three moods, including positive, noncommittal and negative (in some cases, noncommittal is excluded) [3,9,37]. To evaluate the sentiments of user comments, we first propose a 5-point scale of mood. Comments are classified into different scale levels by different score values. Some of the key words for comments are listed in Table 1 [37]. The key words of a mood level are the typical words that represent a group of mood words. Examples corresponding to the key words in Table 1 are provided in Table 2. This table indicates that a score can be obtained for a comment if one mood word is detected in this comment.

Table 1
The look-up table [37] for 5-point score and mood.

Score	Description	Mood	Mood keywords
5	Best	Positive	Surprise, best, extra-, ultra-
4	Better	Positive	Much, very
3	Good	Positive	Happy, good, alive, love, interested, positive, strong
2	Fair	Noncommittal	Open
1	Bad	Negative	Anger, disgust, fear, sadness

Table 2
Selected mood words for different sentiments.

Mood keywords	Words
Open	Understanding, confident, reliable, easy, free, sympathetic
Happy	Great, gay, joyous, lucky, fortunate, delighted, overjoyed, gleeful
Alive	Playful, courageous, energetic, liberated, optimistic, provocative
Good	Calm, peaceful, at ease, comfortable, pleased, encouraged, clever
Love	Loving, considerate, affectionate, sensitive, tender, devoted
Interested	Concerned, affected, fascinated, intrigued, absorbed, inquisitive
Positive	Eager, keen, intent, anxious, inspired, determined, excited
Angry	Irritated, enraged, hostile, insulting, sore, annoyed, upset, hateful
Depressed	Lousy, disappointed, discouraged, ashamed, powerless, diminished
Sad	Tearful, sorrowful, pained, grief, anguish, desolate, desperate

Suppose that there are $x \geqslant 1$ comments posted by UGI viewers. The quality score S can be evaluated by the comment based score in terms of average score utilizing the following equation:

$$S = \begin{cases} 0 & x = 0 \\ \frac{1}{5x}\sum_{i=1}^{x} c_i & x \geqslant 1 \end{cases} \tag{4}$$

where c_i is the comment based score for the i–th comment.

From the discussion in Section 2.1.1, we know that it is very difficult to obtain more comments x when the total number of fans N is fixed. However, it is also very difficult to obtain a higher score S in x comments. Therefore, we use $\tau(S) = 1 - S$ to denote the difficulty of obtaining a high score S.

2.1.3. Analysis of accessibility

Accessibility describes whether this image can be accessed by viewers. For some UGIs, the owner holds the copyright and sometimes limits the accessibility to the public. The UGI is meaningless to a viewer if it is inaccessible. Therefore, accessibility is a binary parameter in our model, which can be described by the following equation:

$$\varphi = \begin{cases} 1 & \textit{obtainable} \\ 0 & \textit{otherwise} \end{cases} \tag{5}$$

2.1.4. Model of presentation measurement

We summarize the above discussions as follows:

- The social circle of a UGI provider can be modeled by a graph model G_i. Based on G_i, the event of a UGI viewer offering his comment is a stochastic event, and the probability that v_i receives at most x comments is $E(x)$. The possibility $E(x)$ is a monotonically decreasing function, which indicates that it is difficult to receive more comments. Therefore, an image may be of higher quality if it receives more comments.
- Comments from viewers can reveal his/her attitude towards the UGI, and the natural words in a comment can be quantified utilizing semantic analysis. It is difficult to receive more positive comments, especially when it is difficult to receive more comments. Therefore, we use $\tau(S)$ to measure the difficulty of receiving a higher score S.
- Only an accessible UGI is meaningful to viewers; therefore, the accessibility is measured by φ and it is included in our model.

We pool the Eqs. (3)–(5) and build the presentation measurement model $\epsilon(x)$ as follows:

$$\epsilon(x) = -\sigma \cdot \varphi \cdot \tau(S) \cdot E(x) \log(\tau(S) \cdot E(x)) \tag{6}$$

where σ is a model parameter.

2.2. Distortion measurement

The intrinsic quality of a UGI measures the perceived image degradation. In most cases, a reference image is needed for quality evaluation and the image is usually an original with perfect quality. However, there is no reference image for a UGI. A UGI is usually captured by the end-user, simply edited, compressed and then published. The degradation of a UGI may result from two aspects. The first aspect is image compression (e.g., quantization losses by JPEG format encoding), and the second is manual re-editing by users. In re-generating the original image, aesthetics and emotions may be involved in re-editing. In this case, it is very difficult to effectively evaluate this type of art [4,13,27,44]. The degradation caused by compression is a type of impairment of signals. No-reference image quality prediction method is an effective approach to predict the perceived quality of these images.

We adopt the no-reference image quality prediction model in [60] in our scheme because this model is suitable for JPEG compressed images and is verified utilizing a LIVE image database [54]. Let μ denote the UGI and $\mu(i,j)$ is the pixel intensity on position (i,j) in μ, the blind image quality $\Psi(\cdot)$ is described by the following equation:

$$\psi(\mu) = \alpha + \beta B(\mu)^{\gamma_1} A(\mu)^{\gamma_2} Z(\mu)^{\gamma_3} \tag{7}$$

where $\alpha, \beta, \gamma_1, \gamma_2$ and γ_3 are model parameters, and

$$B(\cdot) = \frac{B_h(\cdot) + B_v(\cdot)}{2}, \quad A = \frac{A_h(\cdot) + A_v(\cdot)}{2}, \quad Z = \frac{Z_h(\cdot) + Z_v(\cdot)}{2} \tag{8}$$

where

$$\begin{aligned} B_* &= \frac{1}{w(\lfloor h/8 \rfloor - 1)} \sum_{i=1}^{w} \sum_{j=1}^{\lfloor h/8-1 \rfloor} |\mu_*(i, 8j)| \\ A_* &= \frac{1}{7} \left[\frac{8}{w(h-1)} \sum_{i=1}^{w} \sum_{j=1}^{h-1} |\mu_*(i,j)| - B_* \right] \\ Z_* &= \frac{1}{w(h-2)} \sum_{i=1}^{w} \sum_{j=1}^{h-2} z_*(i,j) \end{aligned} \tag{9}$$

where $*$ can be one of h and v, indicating horizontal and vertical pixels, w and h is the width and height of μ, respectively.

2.3. The multi-dimensional preference assessment model

The above discussions suggest that the quality of a UGI is composed of two parts, presentation and distortion measurements. The evaluation of presentation reveals how viewers will favorite the UGI. The distortion measurement indicates the degree to which the UGI preserves the original signals. We describe the model in Eq. (1) utilizing following equation:

$$\varsigma(x, \mu) = f(\epsilon(x), \psi(\mu)) = \omega\epsilon(x) + (1 - \omega)\psi(\mu) \tag{10}$$

where $\omega \in [0, 1]$ is a parameter.

3. Experiments and discussions

3.1. Experiment arrangements

We download UGIs and their corresponding comments from Flickr utilizing keywords such as cat, street, dog, sea and car. The test images are randomly selected from the search results, and the total number is 55. Fig. 3 provides some of the test images. Both subjective and objective experiments were conducted. In subjective experiments, participants are invited to offer their subjective score on quality to each of the test images. The results of the subjective experiments are then used as benchmarks for later comparisons and verifications of the objective experiments.

3.2. Subjective experiments

The subjective experiments followed the strict procedure described in [1], including participant selection, environment and display settings. There are 10 participants (two females and eight males) involved in the subjective experiments who range in age from 21 to 32 with different academic and professional backgrounds. Participants were not involved in similar subjective experiments in the past three months. The single stimulus continuous quality evaluation test method described in [1,34] is applied in our subjective experiments. Each participant is asked to rate their preferences towards a UGI. Although the test procedure for each participant is exactly the same, all of the tests are performed independently to avoid interference.

After the subjective experiments, 550 raw subjective scores were collected. However, these raw data from participants cannot be used as ground truth data directly because noise be present. We further process the raw data to eliminate the noise following an algorithm to improve the confidence of the subjective score.

Suppose that there are P participants and U test UGIs in the subjective experiments, and the subjective score s_{ij} is provided by i-th participant towards j-th UGI. Therefore, we process s_{ij} by Algorithm 1.

Algorithm 1. Process on subjective scores.

Step 1. Calculate the average a_j for each UGI:

$$a_j = \frac{1}{P}\sum_{i=1}^{P} s_{ij} \tag{11}$$

where $j = 1, 2, \ldots, U$.

Step 2. Calculate the standard deviation d_j for each UGI.

$$d_j = \sqrt{\frac{1}{P-1}\sum_{i=1}^{P}(s_{ij} - a_j)^2} \tag{12}$$

where $j = 1, 2, \ldots, U$.

Step 3. Mark s_{ij} as an outlier if $s_{ij} \notin [a_j - d_j, a_j + d_j]$, where $j = 1, 2, \ldots, U$.

Step 4. The i-th participant is marked as outlier if he owns more than 30% outlier s_{ij}, where $i = 1, 2, \ldots, P$. New participant set is determined after all outliers are marked out, and the scores of P^* participants are left.

Step 5. Calculate the ground truth score a_j^* for each UGI:

$$a_j^* = \frac{1}{P^*}\sum_{i=1}^{P^*} s_{ij} \tag{13}$$

where $j = 1, 2, \ldots, U$.

After being processed by Algorithm 1, one participant was identified as an outlier, and all of his submitted scores were eliminated from dataset. Therefore, the final ground truth is calculated based on 495 raw scores obtained from the remaining nine participants.

3.3. Objective experiments

We also use our proposed scheme to measure the objective quality score of the 55 UGIs. These results are compared to the benchmark data obtained to verify the accuracy of quality prediction. The correlation between the predicted score and benchmark data can be used to reveal the performance of our proposed scheme. As seen in Eqs. (10) and (2), the model parameter ω and the probability p is not yet determined. Therefore, discussion of these two parameters is needed. As revealed by the symmetric property of the bi-nominal distribution, the cases are equivalent for the probability p that $P_k|_{p=\alpha} = P_k|_{p=1-\alpha}$, where $0 \leqslant \alpha \leqslant 1$. Therefore, we need to verify the case for $0 \leqslant p \leqslant 0.5$.

These results are summarized in Fig. 4 and Table 3, where different settings for ω and p are selected. The Pearson parameter is widely used to measure the correlation between two datasets, and higher values of the Pearson parameter indicate stronger correlations. Table 3 provides the Pearson parameters for different ω and p. First, we ignore the value of ω and let p vary from 0.1 to 0.5. The Pearson parameter always peaks when $p = 0.1$ regardless the value of ω. This result suggests that people seldom comment on a UGI in social networks. In this case, the UGI should receive a higher quality score if it receives many positive comments. Then, we fix p to 0.1 and vary the value of ω from 0.9 to 0.4. The value of ω is the weight for presentation and distortion measurement. We find that the Pearson parameter peaks when $\omega = 0.6$, which indicates that presentation measurement contributes more to the image quality score. This phenomenon relies on the fact that a UGI is published for self-presentation and thought exchange, and these functionalities have exceeded the traditional usage of images. In addition to the results presented in Table 3, Fig. 4 also indicates the correlations between the predicted quality score and benchmark data when $\omega = 0.6$ and $p = 0.1$. The scatter points are concentrated linearly with some outliers, a pattern that indicates that the prediction effectiveness of our model is high.

Fig. 3. Examples of test UGIs drawn from the Flickr website.

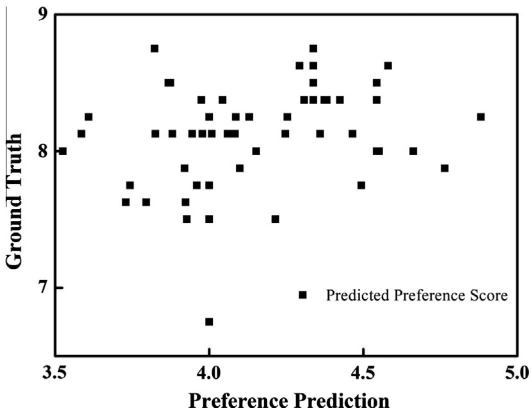

Fig. 4. Correlations between predicted preference scores and benchmark data with $\omega = 0.6$ and $p = 0.1$.

Table 3
The Pearson correlation performance of the predicted preference score compared to the benchmark data, with ω ranging from 0.9 to 0.4 and p is from 0.1 to 0.5.

	$p = 0.1$	$p = 0.2$	$p = 0.3$	$p = 0.4$	$p = 0.5$
$\omega = 0.9$	0.5135	0.4294	0.4487	0.4692	0.4716
$\omega = 0.8$	0.5437	0.4600	0.4753	0.4971	0.5023
$\omega = 0.7$	0.5680	0.4860	0.4961	0.5178	0.5257
$\omega = 0.6$	**0.5779**	0.4998	0.5045	0.5238	0.5337
$\omega = 0.5$	0.5643	0.4939	0.4945	0.5093	0.5193
$\omega = 0.4$	0.5242	0.4659	0.4647	0.4743	0.4824

4. Conclusions

In this paper, we propose a multi-dimensional image quality prediction model for user-generated images in social networks. The model processes the multi-dimensional data from images, texts and social networks utilizing both presentation measurement and distortion measurement. We build a database of images and subjective quality scores, and this database can be used as the benchmark for objective experiments. The experimental results indicate that more weight can be given to presentation measurement (i.e., $\omega = 0.6$) while distortion measurement can be assigned a smaller weight. This result indicates that social functionality is more important than the quality of the image in predicting the quality of user-generated images. Moreover, the Pearson correlation parameter between the predicted score and the ground truth data is 0.5779, which indicates that our model can be utilized to predict image quality in practical environments.

Acknowledgments

This work was supported by Natural Sciences Foundation of China (NSFC) (Grant Nos. 61170194, 61231010 and 61202301).

References

[1] Methodology for the Subjective Assessment of the Quality of Television Pictures, ITU-R BT.500-11.
[2] G. Adomavicius, A. Tuzhilin, Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions, IEEE Trans. Knowl. Data Eng. 17 (6) (2005) 734–749.
[3] S. Baccianella, A. Esuli, F. Sebastiani, Sentiwordnet 3.0: an enhanced lexical resource for sentiment analysis and opinion mining, in: Proceedings of the Seventh Conference on International Language Resources and Evaluation (LREC10), 2010.
[4] S. Bhattacharya, R. Sukthankar, M. Shah, A framework for photo-quality assessment and enhancement based on visual aesthetics, in: Proceedings of the International Conference on Multimedia, ACM, 2010, pp. 271–280.
[5] A. Blum, T. Mitchell, Combining labeled and unlabeled data with co-training, in: Proceedings of the Eleventh Annual Conference on Computational Learning Theory, ACM, 1998, pp. 92–100.
[6] D. Bollegala, N. Okazaki, M. Ishizuka, A preference learning approach to sentence ordering for multi-document summarization, Inf. Sci. 217 (2012) 78–95.
[7] F. Cena, S. Likavec, F. Osborne, Anisotropic propagation of user interests in ontology-based user models, Inf. Sci. 250 (2013) 40–60.
[8] M. Cha, H. Kwak, P. Rodriguez, Y.-Y. Ahn, S. Moon, I tube, you tube, everybody tubes: Analyzing the world's largest user generated content video system, in: Proceedings of the 7th ACM SIGCOMM Conference on Internet Measurement, IMC'07, 2007, pp. 1–14.
[9] H. Chen, D. Zimbra, Ai and opinion mining, IEEE Intell. Syst. 25 (3) (2010) 74–80.
[10] T. Daugherty, M.S. Eastin, L. Bright, Exploring consumer motivations for creating user-generated content, J. Interact. Advertis. 8 (2) (2008) 16–25.
[11] W.J.E. Eleanor, Hyperbole over cyberspace: self-presentation and social boundaries in internet home pages and discourse, Inf. Soc. 13 (4) (1997) 297–327.
[12] G. Erving, The Presentation of Self in Everyday Life, The Overlook Press, New York, 1959. pp. 17–25..
[13] X. Gao, W. Lu, D. Tao, X. Li, Image quality assessment based on multiscale geometric analysis, IEEE Trans. Image Process. 18 (7) (2009) 1409–1423.
[14] Y. Gao, J. Tang, R. Hong, S. Yan, Q. Dai, N. Zhang, T.-S. Chua, Camera constraint-free view-based 3-d object retrieval, IEEE Trans. Image Process. 21 (4) (2012) 2269–2281.
[15] Y. Gao, M. Wang, R. Ji, X. Wu, Q. Dai, 3D object retrieval with hausdorff distance learning, IEEE Trans. Ind. Electron. 61 (4) (2014) 2088–2098.
[16] Y. Gao, M. Wang, R. Ji, Z. Zha, J. Shen, *k*-Partite graph reinforcement and its application in multimedia information retrieval, Inf. Sci. 194 (2012) 224–239.
[17] Y. Gao, M. Wang, D. Tao, R. Ji, Q. Dai, 3-D object retrieval and recognition with hypergraph analysis, IEEE Trans. Image Process. 21 (9) (2012) 4290–4303.
[18] Y. Gao, M. Wang, Z.-J. Zha, J. Shen, X. Li, X. Wu, Visual–textual joint relevance learning for tag-based social image search, IEEE Trans. Image Process. 22 (1) (2013) 363–376.
[19] Y. Gao, M. Wang, Z.-J. Zha, Q. Tian, Q. Dai, N. Zhang, Less is more: efficient 3-d object retrieval with query view selection, IEEE Trans. Multimedia 13 (5) (2011) 1007–1018.
[20] I. Garcia, S. Pajares, L. Sebastia, E. Onaindia, Preference elicitation techniques for group recommender systems, Inf. Sci. 189 (0) (2012) 155–175.
[21] N. Guan, D. Tao, Z. Luo, B. Yuan, Nenmf: an optimal gradient method for non-negative matrix factorization, IEEE Trans. Signal Process. 60 (6) (2012) 2882–2898.
[22] N. Guan, D. Tao, Z. Luo, B. Yuan, Online non-negative matrix factorization with robust stochastic approximation, IEEE Trans. Neural Networks Learn. Syst. 23 (7) (2012) 1087–1099.
[23] T. Holmes, J. Zanker, I like what i see: using eye-movement statistics to detect image preference, J. Vis. 9 (8) (2009). Article id: 385.
[24] J.F. Huete, J.M. Fernández-Luna, L.M. de Campos, M.A. Rueda-Morales, Using past-prediction accuracy in recommender systems, Inf. Sci. 199 (2012) 78–92.
[25] R. Ji, L. Duan, J. Chen, H. Yao, J. Yuan, Y. Rui, W. Gao, Location discriminative vocabulary coding for mobile landmark search, Int. J. Comput. Vis. 96 (3) (2012) 290–314.

[26] R. Ji, Y. Gao, B. Zhong, H. Yao, Q. Tian, Mining flickr landmarks by modeling reconstruction sparsity, ACM Trans. Multimedia Comput. Commun. Appl. 7 (1) (2011) 1–22.
[27] D. Joshi, R. Datta, E. Fedorovskaya, Q.-T. Luong, J. Wang, J. Li, J. Luo, Aesthetics and emotions in images, IEEE Signal Process. Mag. 28 (5) (2011) 94–115.
[28] A.A. Kardan, M. Ebrahimi, A novel approach to hybrid recommendation systems based on association rules mining for content recommendation in asynchronous discussion groups, Inf. Sci. 219 (2013) 93–110.
[29] P. Kazienko, K. Musial, T. Kajdanowicz, Multidimensional social network in the social recommender system, IEEE Trans. Syst. Man Cybernet. – Part A: Syst. Humans 41 (4) (2011) 746–759.
[30] J.M. King, I.A. Elbakri, M. Reed, J. Wrogemann, A clinical image preference study comparing digital tomosynthesis with digital radiography for pediatric spinal imaging, in: Proceedings of SPIE, vol. 7966, 2011, p. 796613.
[31] Y. Koren, R. Bell, C. Volinsky, Matrix factorization techniques for recommender systems, Computer 42 (8) (2009) 30–37.
[32] J. Krumm, N. Davies, C. Narayanaswami, User-generated content, IEEE Pervasive Comput. 7 (4) (2008) 10–11.
[33] D.D. Lee, H.S. Seung, Learning the parts of objects by non-negative matrix factorization, Nature 401 (6755) (1999) 788–791.
[34] J.-S. Lee, F. De Simone, T. Ebrahimi, Subjective quality evaluation via paired comparison: application to scalable video coding, IEEE Trans. Multimedia 13 (5) (2011) 882–893.
[35] T.-W. Lee, Independent Component Analysis, Springer, 1998.
[36] Y.W. Leea, D.M. Strongb, B.K. Kahnc, R.Y. Wang, Aimq: a methodology for information quality assessment, Inf. Manage. 40 (2) (2002) 133–146.
[37] B. Liu, Opinion mining and sentiment analysis, in: M.J. Carey, S. Ceri (Eds.), Web Data Mining, Data-Centric Systems and Applications, Springer, Berlin Heidelberg, 2011, pp. 459–526.
[38] Q. Liu, Y. Yang, X. Wang, L. Cao, Quality assessment on user generated image for mobile search application, in: Advances in Multimedia Modeling, Springer, 2013, pp. 1–11.
[39] W. Liu, D. Tao, Multiview hessian regularization for image annotation, IEEE Trans. Image Process. 22 (7) (2013) 2676–2687.
[40] W. Liu, D. Tao, J. Cheng, Y. Tang, Multiview hessian discriminative sparse coding for image annotation, Comput. Vis. and Image Und. 118 (2014) 50–60.
[41] Y. Luo, D. Tao, B. Geng, C. Xu, S. Maybank, Manifold regularized multitask learning for semi-supervised multilabel image classification, IEEE Trans. Image Process. 22 (2) (2013) 523–536.
[42] A. Maurno, C. Batini, Methodologies for data quality assessment and improvement, ACM Comput. Surveys 41 (3) (2009) 1–52.
[43] W. Meng, L. Hao, T. Dacheng, L. Ke, W. Xindong, Multimodal graph-based reranking for web image search, IEEE Trans. Image Process. 21 (11) (2012) 4649–4661.
[44] A.K. Moorthy, A.C. Bovik, Visual quality assessment algorithms: what does the future hold?, Multimedia Tools Appl 51 (2) (2011) 675–696.
[45] Y. Ni, L. Xie, Z.-Q. Liu, Minimizing the expected complete influence time of a social network, Inf. Sci. 180 (13) (2010) 2514–2527.
[46] L. Parfeni, Flickr Boasts 6 billion Photo Uploads, Tech. rep.
[47] L.L. Pipino, Y.W. Lee, R.Y. Wang, Data quality assessment, Commun. ACM 45 (4) (2002) 211–218.
[48] P. Resnick, H.R. Varian, Recommender systems, Commun. ACM 40 (3) (1997) 56–58.
[49] F. Ricci, L. Rokach, B. Shapira, Introduction to recommender systems handbook, in: F. Ricci, L. Rokach, B. Shapira, P.B. Kantor (Eds.), Recommender Systems Handbook, Springer, US, 2011, pp. 1–35.
[50] H. Richang, T. Jinhui, T. Hung-Khoon, N. Chong-Wah, Y. Shuicheng, C. Tat-Seng, Beyond search: event-driven summarization for web videos, ACM Trans. Multimedia Comput. Commun. Appl. 7 (4) (2011) 35.
[51] H. Richang, W. Meng, L. Guangda, N. Liqiang, C. Zheng-Jun, Zha, Tat-Seng, Multimedia question answering, IEEE Multimedia 19 (4) (2012) 72–78.
[52] H.J. Schau, mary C. Gilly, We are what we post? Self presentation in personal web space, J. Comsum. Res. 30 (3) (2003) 385–404.
[53] G. Seidman, Self-presentation and belonging on facebook: how personality influences social media use and motivations, Person. Indiv. Differ. 54 (3) (2013) 402–407.
[54] H. Sheikh, M. Sabir, A. Bovik, A statistical evaluation of recent full reference image quality assessment algorithms, IEEE Trans. Image Process. 15 (11) (2006) 3440–3451.
[55] H. Shin, J. Lee, Impact and degree of user sociability in social media, Inf. Sci. 196 (2012) 28–46.
[56] A. Susarla, J.-H. Oh, Y. Tan, Social networks and the diffusion of user-generated content: evidence from youtube, Inf. Syst. Res. 23 (1) (2012) 23–41.
[57] D. Tao, L. Jin, Discriminative information preservation for face recognition, Neurocomputing 91 (2012) 11–20.
[58] D. Tao, L. Jin, W. Liu, X. Li, Hessian regularized support vector machines for mobile image annotation on the cloud, IEEE Trans. Multimedia 15 (4) (2013) 833–844.
[59] R.Y. Wang, D.M. Strong, Beyond accuracy: what data quality means to data consumers, J. Manage. Inf. Syst. 12 (4) (1996) 5–33.
[60] Z. Wang, H. Sheikh, A. Bovik, No-reference perceptual quality assessment of jpeg compressed images, in: 2002 International Conference on Image Processing, vol. 1, 2002, pp. I-477– I-480.
[61] Z. Yuan, E. Oja, A fastica algorithm for non-negative independent component analysis, in: Independent Component Analysis and Blind Signal Separation, Springer, 2004, pp. 1–8.
[62] F. Zhuang, G. Karypis, X. Ning, Q. He, Z. Shi, Multi-view learning via probabilistic latent semantic analysis, Inf. Sci. 199 (2012) 20–30.
[63] S. Zuffi, C. Brambilla, R. Eschbach, A. Rizzi, Comparing image preference in controlled and uncontrolled viewing conditions, J. Electron. Imag. 19 (4) (2010) 043014.