

Image Quality Assessment by Combining Masking Texture and Perceptual Color Difference Model

Zhisen Tang¹, Yuanlin Zheng^{1,2*}, Wei Wang¹, and Kaiyang Liao^{1,3}

¹ Faculty of Printing, Packaging Engineering and Digital Media Technology, Xi'an University of Technology, China

[e-mail: zhisentang@163.com, zhengyuanlin@xaut.edu.cn, wangweixaut@foxmail.com, liaokaiyang@xaut.edu.cn]

² Key Lab of Printing and Packaging Engineering of Shaanxi Province, China
[e-mail: zhengyuanlin@xaut.edu.cn]

³ Printing and Packaging Engineering Technology Research Centre of Shaanxi Province, China
[e-mail: liaokaiyang@xaut.edu.cn]

*Corresponding author: Yuanlin Zheng

*Received March 3, 2019; revised May 1, 2019; accepted March 23, 2020;
published July 31, 2020*

Abstract

Objective image quality assessment (IQA) models have been developed by effective features to imitate the characteristics of human visual system (HVS). Actually, HVS is extremely sensitive to color degradation and complex texture changes. In this paper, we firstly reveal that many existing full reference image quality assessment (FR-IQA) methods can hardly measure the image quality with contrast and masking texture changes. To solve this problem, considering texture masking effect, we proposed a novel FR-IQA method, called Texture and Color Quality Index (TCQI). The proposed method considers both in the masking effect texture and color visual perceptual threshold, which adopts three kinds of features to reflect masking texture, color difference and structural information. Furthermore, random forest (RF) is used to address the drawbacks of existing pooling technologies. Compared with other traditional learning-based tools (support vector regression and neural network), RF can achieve the better prediction performance. Experiments conducted on five large-scale databases demonstrate that our approach is highly consistent with subjective perception, outperforms twelve the state-of-the-art IQA models in terms of prediction accuracy and keeps a moderate computational complexity. The cross database validation also validates our approach achieves the ability to maintain high robustness.

Keywords: Image quality assessment, full-reference, masking texture, color difference, human visual system

This work was supported in part by the National Natural Science Foundation of China (NO. 61671376) and in part by the Scientific Research Project of Shaanxi Provincial Department of Education (NO. 18JK0556).

1. Introduction

The past decades have witnessed the tremendous development of the digital media and communication technologies. Nowadays, people put forward higher demand for visual quality. Thus, quantifying the degree of distortion in image is the essence for most applications [1], including image digitization, transmission and compression. Because the subjective assessment lies largely in the individual physiological and psychological factors, the deployment of objective image quality assessment (IQA) which measures the visual quality automatically is particularly important. According to the degree of the availability of pristine image information, objective IQA models can typically be classified into three types: full reference IQA (FR-IQA), reduced reference IQA (RR-IQA) and no reference IQA (NR-IQA) models. In this paper, we concentrate our attention on FR-IQA, because it includes an extensive real applications, such as optimization of coding technologies [2] and weak supervision for learning NR models [3, 4].

The conventional methods, i.e., peak signal-to-noise ratio (PSNR) and mean squared errors (MSE), have the lower prediction results since they do not highly correlate with subjective perception. Therefore, the mass of IQA models based on the human visual system (HVS) have been deployed recently. Generally, FR-IQA methods evaluate image quality by measuring the similarity or difference between the distorted and the original images, and utilize the pooling technology to convert the similarity into a quality score. Structural similarity index (SSIM) [5] is deemed to be a landmark of FR-IQA methods, which extracts the luminance, contrast and structure features and employs the average pooling to generate a single quality score. Because of the great success of SSIM, its extended versions, called MS-SSIM [6] and IW-SSIM [7], have been introduced by multiscale and information content weighted SSIM. Moreover, the gradient magnitude (GM) which can effectively capture the local structural information is widely applied in IQA studies [8-10]. FSIM [9] introduces the gradient magnitude (GM) and phase congruency (PC) as the complementary features to capture the image quality. Similarity, in VSI [10], the GM and visual saliency (VS) are utilized to reflect the degree of distortion in image. Although these methods in low-level features for IQA achieve better performance than PSNR and MSE, they are not in compliance with subjective mean opinion scores (MOS) and without considering color degradation. Gradient magnitude similarity deviation (GMSD) [11] and mean deviation similarity index (MDSI) [12] adopt the GM and deviation to calculate the final quality score. Although GMSD and MDSI are run faster than other models, their prediction accuracies need to be further improved.

Furthermore, some other methods simulate the visual processing progress in HVS. Their performances mainly reflected the many visual characteristics of HVS, such as visual masking, nonlinear perception and contrast sensitivity function (CSF) and bandpass property. In order to effectively extract the visual features, the contrast masking (CM) effect [13] and color visual perception threshold (VPT) [14, 15] have been frequently applied in some FR-IQA methods. The CM effect reveals that HVS has different sensitivities to distortion factors depending on background luminance and texture characteristics in image. The perception of HVS is not only related to the average brightness of the neighboring region, but also to the spatial changes of the brightness of the neighboring region. Perceptual characteristic theory considers that color perception effect is caused by the non-uniform topological mapping of the color space on the perceptive neurons. Meanwhile, as their spatial distance increases, the ability to distinguish difference between two colors becomes easier. And Modern physiology

researches [16] show that HVS cannot distinguish the colors with smaller color difference.

In the approaches above, although their prediction accuracies are not bad, they cannot maintain the good performances across all databases. Furthermore, SSIM considers different distortions have same importance to visual perception. Actually, the different distortions have different contributions to the image quality. Other IQA models avoid the drawback by using a weighted summation, but there is no universal method available for appropriate the weighting coefficients, which need to be determined in weighted summation. Moreover, [8] and [27] have researched the simple summation or multiplication may lead to the relationship between the distortions and image quality score to be linear in the methods of above-mentioned. Fortunately, the pooling strategies by using machine learning can overcome the drawbacks, which introduce the human vision knowledge to the trained model. Thus, the objective quality predicted by machine learning model will maintain higher consistency with subjective perception.

In order to overcome the limitations above, we proposed a FR-IQA method for color image, called Texture and Color Quality Index (TCQI), which provides an accurate prediction of the input images. We extract the the masking texture maps from the distorted images and their corresponding reference images, which are motivated by the fact that human perceptual masking effect is affected by the background illumination and texture complexity. Afterwards, we analyze the color perception mechanism and extract color difference by using CIE $L^*a^*b^*$ [17] (details can be found in Section 2.1). Finally, we replace the formula-based with learning-based to reflect the relationships between the distortions and subjective mean opinion score (MOS).

This paper makes two contributions:

1) Considering the texture masking effect, we introduce a novel feature extraction method to extract the masking texture map. Because the background illumination affects the texture feature, we employ a function to fit the relationship between background illumination and texture. Furthermore, the texture is also affected by self-complexity, and we adopt the max value of four Laws' texture features to represent the complex texture information. The experiments show that the proposed masking texture feature is quality-aware and it can be applied to other image quality assessment algorithms.

2) We found that traditional color IQA metrics which compute the similarity by using YIQ or LMN color space, always fail in evaluating the distorted image with contrast change. Compared with the advantages and disadvantages of different color spaces, the CIE $L^*a^*b^*$ [17] is more appropriate way to measure color degradation. In addition, we consider the HVS cannot distinguish the colors with smaller color difference, and a novel color difference feature is proposed to capture the color degradation between reference and distorted image. Experimental results verify that the proposed color difference approach outperforms other color measurements. Furthermore, the color difference feature can also be used in other FR-IQA models especially in learning-based methods for color distortion evaluation.

The rest of this paper is organized as follows. Section 2 briefly retrospects the color spaces, color difference computations and machine learning tools in existing FR-IQA. The details proposed IQA algorithm is depicted in Section 3. Section 4 gives experimental results and analysis. Finally, conclusions are shown in Section 5.

2. Related Work

The color perception of HVS cannot be reflected if only grayscale features are used. Therefore, it is necessary to adopt a method to color image quality assessment. In the process of IQA,

most digital images exist in the form of RGB, which are converted into other color spaces, such as YIQ [18] and LMN [18]. [16] lists the representative FR-IQA for color image and analyze the color space. The different color spaces have the different color characteristics, which may cause different results. [16] has proved that YIQ and LMN color spaces are obtained by simple linear transform from RGB color space, which are not well aligned with HVS. Furthermore, CIE $L^*a^*b^*$ computes the color difference in perceptually uniform space, which is appropriate for defining a simple yet precise measure of color difference and adopts a contrast sensitivity function (CSF) filtering to imitate spatial sensitivities of HVS. We found that the simple color similarity in YIQ and LMN color spaces always fail to be consistent with HVS when evaluating the distorted images with contrast change. As shown in Fig. 1 that an example of this situation. Fig. 1 (a) is a reference image in TID2013 database, and Fig. 1(b-f) are its five contrast change distorted image with different distortion level. We list two typical methods, i.e., FSIMc [9] and VSI [10], they measure the color degradation in YIQ and LMN color spaces, respectively. CSFSIM and CSVSI denote the indicators of color measurement in FSIM and VSI, respectively. As the distortion level increases, the MOS does not continue to decline. But the similarity values computed by FSIMc, VSI, CSFSIM and CSVSI decrease gradually. For example, Fig. 1 (d-e), the MOS increases from 5.20000 to 6.57500, but all of the four similarity indicators decrease. It shows that these methods are not consistent with human subjective perception.

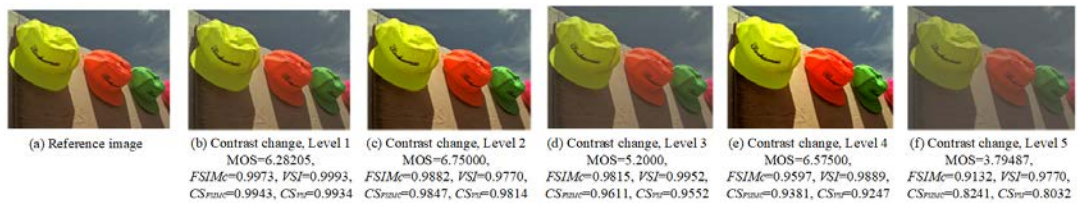


Fig. 1. Examples of distorted images with contrast change on TID2013. (a) Reference image, (b-f) contrast change images with five different distortion levels.

Modern physiology researches show that there are three different color receptors in the retina. They are three kinds of chromatic pyramidal cells, and each of which has different spectral sensitivity characteristics. In the nervous system, there are three types of reactions which contain luminance reaction, red-green and yellow-blue reactions. Fortunately, the CIE $L^*a^*b^*$ color space is widely used in the image processing, printing and dyeing industries, which includes three components, i.e., luminance L^* and chromaticity a^* and b^* . The components a^* and b^* represent the ranges from red to green and yellow to blue, respectively. The characteristic L^* is compliance with the luminance reaction in the nervous system. The components a^* and b^* of the CIE $L^*a^*b^*$ color space can denote the red-green and yellow-blue reactions of ganglion cells, respectively. Compared with other color model, CIE $L^*a^*b^*$ is a device-independent model which has a larger gamut. Hence, the CIE $L^*a^*b^*$ model is used to decompose an image into luminance and color channels in this paper.

Generally, the existing FR-IQA methods consist of two steps. First, they extract features and compute the similarity maps. The second stage is to fuse the similarity features and calculate the single quality score. The average and the weighting pooling strategies are used in [5, 6]. Other pooling techniques, such as information content pooling [7, 19], saliency pooling [10, 20] and percentile pooling [21] are also utilized in FR-IQA models. However, these simple pooling strategies may cause the relationship between the distortion factors and image quality score to be linear. Fortunately, the machine learning technique is able to overcome these limitations, which can simulate the visual weighting function in IQA.

The learning-based methods are extensively applied in the field of IQA, including neural networks (NN) [22, 23], support vector regression (SVR) [24-26], extreme learning machine (ELM) [27] and random forest (RF) [8, 28, 29]. In [24], multiple features are employed for complementary representation of image quality and SVR is used as a regression tool. Moreover, [27] employs ELM as learning model in FR-IQA, which achieves faster learning speed. [29] uses RF to regress the similarity features obtained from Difference of Gaussian (DOG) frequency bands to generate the final quality score. A deep neural network-based FR-IQA model is proposed in [30], which is trained by ten convolutional layers and five pooling layers in step of features extraction, and two fully connected layers act a regression function. Although the deep neural network-based methods show good performance, they are too complicated, time-consuming in training step, and contain the larger number of parameters. Therefore, we replace the formula-based with RF to reflect the relationships between the distortions and subjective mean opinion score (MOS). Compared with other learning-based models, RF has only two parameters, and they are set by default. Meanwhile, RF exhibits the outstanding ability in terms of prediction accuracy [8]. Specifically, we show that RF outperforms SVR and NN both in prediction accuracy and robustness. Details can be found in Section 4.

3. Proposed IQA Model

In this section, a novel FR-IQA method by combining masking texture and perceptual color difference model is proposed, called texture and color quality index (TCQI). The overview of the proposed method TCQI is illustrated in Fig. 2. Before the extraction masking texture, GM and color difference, the RGB images are converted into CIE $L^*a^*b^*$ color space. To maintain the low complexity of the proposed algorithm, only three features extracted from luminance L^* channel, i.e., masking texture, GM and gradient orientation (GO) features. Afterwards, the color difference is calculated in L^* , a^* and b^* channels. It is noteworthy that the masking texture feature and color difference are limited by experience threshold. Finally, these features and MOS are fused with RF to predict the final quality score.

3.1 Color Space Conversion

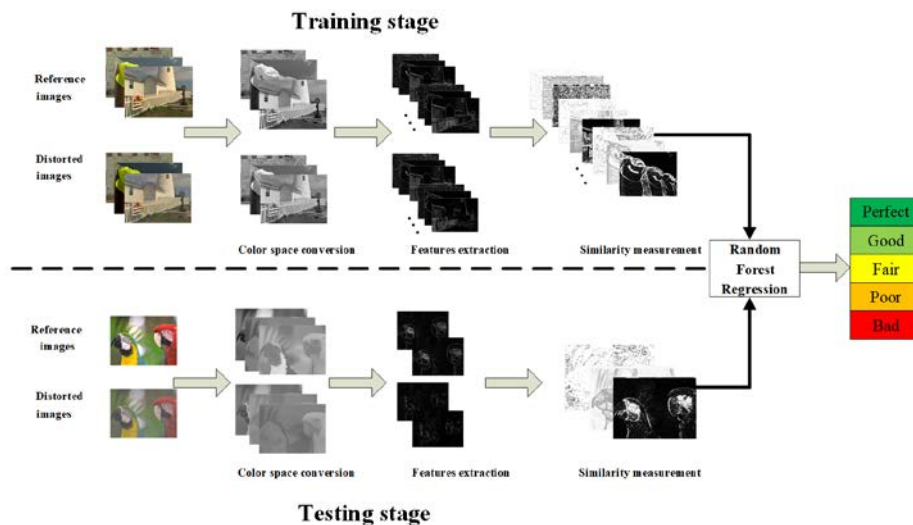


Fig. 2. The flow chart of the proposed TCQI.

In this paper, the input images are converted into CIE L*a*b* color space to decorrelate the luminance L* and chrominance a* and b* channels. According to the international standard [17], the input RGB color image needs to be converted into CIE XYZ and then output the CIE L*a*b* by a nonlinear function. The matlab function 'makecform' is adopted to convert color space.

The original RGB values are normalized to [0, 1] by gamma function. Then, the CIE XYZ is defined as:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.4124 & 0.3576 & 0.1805 \\ 0.2126 & 0.7152 & 0.0722 \\ 0.0193 & 0.1192 & 0.9505 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}, \quad (1)$$

the CIE L*a*b* color space is computed as:

$$\begin{cases} L^* = 116f(Y/Y_0) - 16 \\ a^* = 500[f(X/X_0) - f(Y/Y_0)], \\ b^* = 200[f(Y/Y_0) - f(Z/Z_0)] \end{cases} \quad (2)$$

where X_0 , Y_0 and Z_0 are reference display white coordinates of illuminant D_{65} with $X_0=0.9505$, $Y_0=1.0000$ and $Z_0=1.0890$, and the $f(t)$ is computed as:

$$f(t) = \begin{cases} t^{1/3} & t > 0.008856 \\ 7.877t + 16/116 & t \leq 0.008856 \end{cases}, \quad (3)$$

3.2 The Masking Texture Feature Extraction And Dissimilarity Computation

Visual masking effect is an important characteristic of HVS, which denotes one existing stimulus makes another stimulus invisible. The visual masking effect is a localized effect, which can be affected by the background illumination and texture complexity. Therefore, considering the influence of background illumination, we decide to introduce a function to imitate the texture information. At first, the local average background luminance is computed, and we extract the four kinds of Laws' texture. Then, the max value of them is used to represent complex texture information. Finally, a function is employed to fit the relationship between background illumination and complex texture information, and it is adopted to obtain the final masking texture feature. The process of extraction of the masking texture feature is as follows:

for an input image f , the local average background luminance is calculated as:

$$bg(i, j) = \frac{1}{32} \left(\sum_{x=i-1}^{i+1} \sum_{y=j-1}^{j+1} f(x, y) + \sum_{x=i-2}^{i+2} \sum_{y=j-2}^{j+1} f(x, y) \right), \quad (4)$$

Laws' texture measurement is a commonly methods for texture feature extraction. The 1-D masks of Law' texture filters include E_5 , L_5 , S_5 , W_5 and R_5 , which denote the different local textural information of the edge, level, spot, wave and ripple texture of an image. The five 1-D masks are $E_5=[-1, -2, 0, 2, 1]$, $L_5=[1, 4, 6, 4, 1]$, $S_5=[-1, 0, 2, 0, -1]$, $W_5=[-1, 2, 0, -2, 1]$ and $R_5=[1, -4, 6, -4, 1]$. 2-D filter operators can be generated by convolution operation of two 1-D filter masks. Fig. 3 shows the four 2-D Law's filters, including E_5L_5 , L_5E_5 , S_5L_5 and L_5S_5 , which are used to extract the vertical edge, horizontal edge, vertical spot and horizontal spot, respectively. In this paper, the vertical edge, horizontal edge, vertical spot and horizontal spot are defined as:

$$\begin{aligned}
te_{E_5L_5} &= E_5L_5 * f(x), \\
te_{L_5E_5} &= L_5E_5 * f(x), \\
te_{S_5L_5} &= S_5L_5 * f(x), \\
te_{L_5S_5} &= L_5S_5 * f(x),
\end{aligned} \tag{5}$$

where $te_{E_5L_5}$, $te_{L_5E_5}$, $te_{S_5L_5}$ and $te_{L_5S_5}$ are the vertical edge, horizontal edge, vertical spot and horizontal spot, respectively. “*” is the convolution operation. The $f(x)$ denotes the input image. After the convolution operation, the max value of texture is defined as:

$$te_{\max} = \max(te_{E_5L_5}, te_{L_5E_5}, te_{S_5L_5}, te_{L_5S_5}), \tag{6}$$

then, the masking texture feature is calculated as:

$$mte = \alpha \cdot te_{\max} + \beta, \tag{7}$$

where α and β are two functions relative to the local average background luminance. α and β are computed as:

$$\alpha(i, j) = k_1 \cdot bg(i, j) + k_2, \tag{8}$$

$$\beta(i, j) = k_3 - k_4 \cdot bg(i, j), \tag{9}$$

In this paper, denote by mte_r and mte_d the two masking texture features extracted from reference and distorted images, respectively. The texture dissimilarity indicator between reference and distorted images is defined as:

$$S_{mte} = \frac{2 \cdot mte_r \cdot mte_d + C_1}{mte_r^2 + mte_d^2 + C_1}, \tag{10}$$

where C_1 is a positive constant to avoid the fraction instability.

Fig. 4 shows the reference (distorted) image and its masking texture map. It is obvious that the masking texture map changes when the image is affected by distortion.

$$\begin{array}{cccc}
\begin{bmatrix} -1 & -4 & -6 & -4 & -1 \\ -2 & -8 & -12 & -8 & -2 \\ 0 & 0 & 0 & 0 & 0 \\ 2 & 8 & 12 & 8 & 2 \\ 1 & 4 & 6 & 4 & 1 \end{bmatrix} &
\begin{bmatrix} -1 & -2 & 0 & 2 & 1 \\ -4 & -8 & 0 & 12 & 4 \\ -6 & -12 & 0 & 12 & 6 \\ -4 & -8 & 0 & 8 & 4 \\ -1 & -2 & 0 & 2 & 1 \end{bmatrix} &
\begin{bmatrix} -1 & -4 & -6 & -4 & -1 \\ 0 & 0 & 0 & 0 & 0 \\ 2 & 8 & 12 & 8 & 2 \\ 0 & 0 & 0 & 0 & 0 \\ -1 & -4 & -6 & -4 & -1 \end{bmatrix} &
\begin{bmatrix} -1 & 0 & 2 & 0 & -1 \\ -4 & 0 & 8 & 0 & -4 \\ -6 & 0 & 12 & 0 & -6 \\ -4 & 0 & 8 & 0 & -4 \\ -1 & 0 & 2 & 0 & -1 \end{bmatrix} \\
\text{(a)}E_5L_5 & \text{(b)}L_5E_5 & \text{(c)}S_5L_5 & \text{(d)}L_5S_5
\end{array}$$

Fig. 3. Laws' filter masks

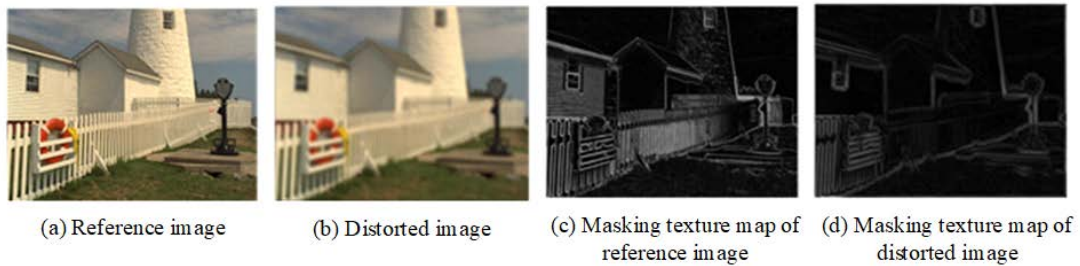


Fig. 4. Illustration of reference (distorted) image and its masking texture map. (a) Reference image, (b) distorted image and their respective masking texture map (c) and (d).

3.3 Chromatic Feature Extraction And Similarity Measure

The chrominance information is also an indispensable factor for HVS understands an image.

In this paper, we adopt the Euclidean distance in CIE L*a*b* color space to measure the color distortion:

$$\Delta E = \sqrt{(L_r^* - L_d^*)^2 + (a_r^* - a_d^*)^2 + (b_r^* - b_d^*)^2}, \quad (11)$$

where L_r^* (L_d^*), a_r^* (a_d^*) and b_r^* (b_d^*) indicate three channels of reference (distorted) image in CIE L*a*b* color space, respectively. And they are computed by Eq. (1)-(3). Because HVS cannot distinguish the colors with smaller color difference [17], an experience threshold function is adopted to simulate the characteristic:

$$\Delta E(i, j) = \begin{cases} 0 & , \quad \Delta E(i, j) < \text{Threshold} \\ \Delta E(i, j) & , \quad \Delta E(i, j) > \text{Threshold} \end{cases}, \quad (12)$$

the average value and standard deviation of ΔE is calculated as:

$$\Delta \bar{E} = \sqrt{\frac{1}{mn} \sum_{i=1}^m \sum_{j=1}^n \Delta E(i, j)}, \quad (13)$$

$$\sigma_{\Delta \bar{E}} = \sqrt{\frac{1}{mn} \sum_{i=1}^m \sum_{j=1}^n (\Delta E(i, j) - \Delta \bar{E})^2}, \quad (14)$$

3.4 The Gradient Feature Extraction And Similarity Computation

Because the HVS is sensitive to structural information [31], the gradient feature has been widely employed in the field of IQA [8-11]. Gradient feature can reflect the intensity and contrast changes in image, and it has the ability to extract the structural information. The GM feature can be obtained by the vertical and horizontal components which are computed by convolving the image with an edge operator. In this paper, we use the Scharr [32] operator since it keeps the lowest computational cost. The vertical and horizontal components of GM are computed as:

$$G_y(x) = \frac{1}{16} \begin{bmatrix} 3 & 10 & 3 \\ 0 & 0 & 0 \\ -3 & -10 & -3 \end{bmatrix} * f(x), \quad G_x(x) = \frac{1}{16} \begin{bmatrix} 3 & 0 & -3 \\ 10 & 0 & -10 \\ 3 & 0 & -3 \end{bmatrix} * f(x), \quad (15)$$

where $G_y(x)$ and $G_x(x)$ represent the vertical and horizontal GM maps, respectively. “*” is the convolution operation. The $f(x)$ denotes the input image. Thus, the GM and gradient orientation (GO) are defined as:

$$G(x) = \sqrt{G_y^2(x) + G_x^2(x)}, \quad (16)$$

$$\theta(x) = \arctan\left(\frac{G_y(x)}{G_x(x)}\right), \quad (17)$$

where $G(x)$ and $\theta(x)$ indicate the GM and gradient orientation (GO) maps, respectively.

Similar to [24], we adopt the chi-square distance to calculate the GM similarity between the reference image I_r and distorted image I_d

$$\chi_{gm}^2(I_r, I_d) = \frac{1}{mn} \sum_{x=1}^{mn} \frac{(I_r(x) - I_d(x))^2}{I_r(x) + I_d(x)}, \quad (18)$$

where $[m, n]$ is the size of the image, χ_{gm}^2 denotes the GM similarity.

Moreover, the gradient orientation similarity is calculated as:

$$S_{or}(x) = \frac{2\theta_r(x)\theta_d(x) + C_2}{\theta_r^2(x) + \theta_d^2(x) + C_2}, \quad (19)$$

where the θ_r and θ_d denote the gradient orientation of I_r and I_d , respectively. C_2 is the positive invariable to control the numerical stability.

3.5 Regression Tool

There are many regression methods, e.g., RF, SVR and NN, they have been widely used in existing IQA models [8, 22, 24, 29]. In this paper, we employ RF to build a mathematical function to reflect the relationship between quality score and distortion factors since the regression performance of RF evidently outperforms NN and SVR. Furthermore, RF has been applied in hardware for real-time application. This is why we choose RF as regression in this paper. The comparisons of RF, SVR and NN are shown in Section 4. After the features extraction, we can acquire a 6-D feature vector, it can be indicated by

$$q = f\{\overline{S_{mte}}, \sigma_{S_{mte}}, \overline{\Delta E}, \sigma_{\Delta E}, \chi_{gm}^2, \overline{S_{or}}\}, \quad (20)$$

where the six features have already described in Sections 3.1-3.3, and they are summarized in **Table 1**. A set of 6-D feature vectors and subjective scores are used to construct a regression model in training step. Afterwards, the feature vectors obtained from the testing dataset are directly mapped to the objective scores by the learned model. Furthermore, the parameters of the proposed TCQI method are exhibited in **Table 2**.

Table 1. The six extracted features in proposed method

Feature Type	Symbol	Feature Description	Equation
Texture	$\overline{S_{mte}}$	Mean value of masking texture similarity map	(4)-(10)
	$\sigma_{S_{mte}}$	Standard deviation of masking texture similarity map	
Color	$\overline{\Delta E}$	Mean value of color aberration	(11)-(14)
	$\sigma_{\Delta E}$	Standard deviation of color aberration	
Gradient	χ_{gm}^2	χ^2 distance between two gradient magnitude maps	(15)-(19)
	$\overline{S_{or}}$	Mean value of gradient orientation similarity map	

Table 2. The parameters setting in the proposed TCQI

TCQI	Parameter	Value
Texture	K_1	0.0001
	K_2	0.115
	K_3	0.5
	K_4	0.01
	C_1	0.01
Color	Threshold	2
Gradient	C_2	100
RF regression	Ntree	500
	Mtry	2

4. Experimental Classification Results and Analysis

4.1 Experimental Databases

Experiments are conducted on five large-scale IQA databases, including TID2013 [33], TID2008 [34], CSIQ [35], LIVE [36], and CCID2014 [37]. The main information of these databases is listed in **Table 3**.

Table 3. Information about five public databases

Databases	Reference images	Distorted images	Distortion type	Resolution
LIVE	29	779	5	Various
TID2008	25	1700	17	512×384
TID2013	25	3000	24	512×384
CSIQ	30	866	6	512×512
CCID2014	15	655	5	Various

To accurately evaluate the prediction ability of IQA model, we adopt four popular performance metrics, i.e., Pearson linear correlation coefficient (PLCC), Spearman rank-order coefficient (SROCC), Kendall rank-order correlation coefficient (KROCC) and root mean squared error (RMSE). Moreover, we utilize a nonlinear logistic regression recommended by the Video Quality Experts Group (VQEG) [38], which provides a better fit for all data and lead the predicted scores on the same scale as the MOS. The nonlinear logistic function is defined as:

$$y = \beta_1 \left(0.5 - \frac{1}{1 + e^{\beta_2(x - \beta_3)}} \right) + \beta_4 x + \beta_5, \quad (21)$$

where x denote the objective quality score, and y represents mapped out. $\beta_i (i \in [1, 5])$ are free parameters to be fitted by minimizing the sum of squared difference y and MOS. These five parameters are estimated by using MATLAB function `nlinfit`. More details about explanations of these five parameters can be found in [38].

4.2 Cross Validation

Because the influence of the image content in quality prediction is significant, it may affect the accuracy of the prediction. In order to ensure that the image contents in presented training subset do not appear in testing subset, all images in each database are not randomly divided into two subsets (80% for training and remaining 20% are used for testing). As shown in [8, 27], we adopt the k -fold to effectively avoid the situation of the over/under-fitting in the learning-based models. All the images in database are divided into two subsets (training and testing subsets) by image content information. Note that we divide all distorted images into k disjoint subsets according to image content information, and the number of samples in each subset is equal or roughly equal. For one subset is utilized testing, the remaining subsets act as training. The final prediction is obtained by the average result of the k testing trials. Furthermore, to achieve the convincing result, the processes of train-test are repeated 1000 times, and the median is used as testing result. To this end, the value of k is similar to [8, 27], TID2013 and TID2008 databases are classified into eight subsets according to the different image contents, respectively. Similarly, LIVE and CSIQ databases are divided into ten subsets, respectively. CCID 2014 database is divided into seven subsets.

4.3 Overall Performance Evaluation

In order to demonstrate the prediction accuracy of proposed method TCQI, we compare 12 state-of-the-art FR-IQA methods, i.e. SSIM [5], IW-SSIM [7], VSI [10], FSIM [9], GSM [39], IFS [40], DSCSI [16], GMSD [11], DOG-SSIMc [29], MDSI [12], PSIM [21], and SCQI [41] with TCQI. Table 4 lists the PLCC SROCC, KROCC and RMSE of TCQI and other metrics on five public image databases. It is noticed that the best results is highlighted in red boldface. It can be seen in Table 4 that the prediction accuracy of TCQI precedes all other metrics by a large margin on five databases. By contrast, there is no method that can attain the best results

in all databases. This means that TCQI extracts the masking texture and color difference can effectively reflect the image quality perception of HVS. At the same time, the **Table 4** shows the comprehensive performance which is computed by taking weighted average (W. A). Specifically, the W. A depends on the numbers of the distorted images. The W. A also proves the proposed method performs better than other metrics. Note that **Table 4** does not list the the W. A value of RMSE, since the RMSE scope differently in five databases. Moreover, it should be noted that TCQI has the best prediction performance in contract distortion images, i.e., CCID2014 database.

Table 4. Performance comparison of 13 IQA models on five databases

		SSIM	IW-SSIM	FSIMc	GSM	GMSD	IFS	VSI	DSCSI	PSIM	MDSI	SCQI	DOG-SSIMc	TCQI
TID 2013	PLCC	0.7895	0.8319	0.8769	0.8464	0.8590	0.8791	0.9000	0.8782	0.9080	0.9085	0.9071	0.9284	0.9448
	SROCC	0.7417	0.7779	0.8510	0.7946	0.8044	0.8697	0.8965	0.8744	0.8925	0.8899	0.9052	0.9143	0.9339
	KROCC	0.5588	0.5977	0.6665	0.6255	0.6339	0.6785	0.7183	0.6862	0.7161	0.7123	0.7327	0.7542	0.7821
	RMSE	0.7608	0.6880	0.5959	0.6603	0.6346	0.5905	0.5404	0.5930	0.5193	0.5181	0.5219	0.4606	0.4030
TID 2008	PLCC	0.7732	0.8579	0.8762	0.8422	0.8788	0.8810	0.8762	0.8445	0.9078	0.9160	0.8899	0.9270	0.9428
	SROCC	0.7749	0.8559	0.8840	0.8504	0.8907	0.8903	0.8979	0.8634	0.9120	0.9208	0.9051	0.9235	0.9392
	KROCC	0.5768	0.6636	0.6991	0.6596	0.7092	0.7009	0.7123	0.6651	0.7395	0.7515	0.7294	0.7721	0.7966
	RMSE	0.8511	0.6895	0.6468	0.7235	0.6404	0.6349	0.6466	0.7187	0.5628	0.5383	0.6120	0.5036	0.4426
CSIQ	PLCC	0.8613	0.9114	0.9192	0.8964	0.9541	0.9576	0.9279	0.9313	0.9642	0.9531	0.9268	0.9445	0.9766
	SROCC	0.8756	0.9213	0.9310	0.9108	0.9570	0.9581	0.9423	0.9417	0.9621	0.9569	0.9434	0.9488	0.9738
	KROCC	0.6907	0.7529	0.7690	0.7374	0.8129	0.8158	0.7857	0.7787	0.8273	0.8130	0.7870	0.8128	0.8643
	RMSE	0.1334	0.1063	0.1034	0.1164	0.0786	0.0757	0.0979	0.0956	0.0696	0.0795	0.0986	0.0857	0.0532
LIVE	PLCC	0.9449	0.9522	0.9613	0.9512	0.9603	0.9586	0.9482	0.9434	0.9584	0.9659	0.9344	0.9515	0.9748
	SROCC	0.9479	0.9567	0.9645	0.9561	0.9603	0.9599	0.9524	0.9487	0.9622	0.9667	0.9406	0.9455	0.9721
	KROCC	0.7963	0.8175	0.8363	0.8150	0.8268	0.8254	0.8058	0.7982	0.8294	0.8395	0.7835	0.7971	0.8645
	RMSE	8.9455	8.3472	7.5296	8.4327	7.6214	7.7764	8.6817	9.0635	7.7996	7.0790	9.7355	4.9045	4.8723
CCID 2014	PLCC	0.8308	0.8342	0.8204	0.8073	0.8521	0.7936	0.8209	0.7586	0.8386	0.8576	0.8200	0.8612	0.9021
	SROCC	0.8174	0.7811	0.7657	0.7768	0.8077	0.7227	0.7734	0.7400	0.8003	0.8128	0.7811	0.8270	0.8718
	KROCC	0.6106	0.5898	0.5707	0.5711	0.6100	0.5337	0.5735	0.5468	0.6038	0.6181	0.5812	0.6398	0.6912
	RMSE	0.3640	0.3606	0.3739	0.3859	0.3422	0.3978	0.3734	0.4260	0.3562	0.3363	0.3734	0.3318	0.2769
W. A	PLCC	0.8156	0.8617	0.8861	0.8596	0.8862	0.8901	0.8956	0.8726	0.9140	0.9175	0.9002	0.9263	0.9476
	SROCC	0.7964	0.8348	0.8736	0.8388	0.8619	0.8819	0.8972	0.8757	0.9050	0.9070	0.9022	0.9161	0.9386
	KROCC	0.6108	0.6566	0.6970	0.6636	0.6936	0.7037	0.7214	0.6919	0.7376	0.7396	0.7301	0.7599	0.7965

To further test the superiority of TCQI, we compare the performance of TCQI with four representative machine learning models, i.e., LLM [22], [24], CF-MMF [26] and DOG-SSIMc [29]. To be fair, the training-test processes of all learning-based methods are the same as TCQI. **Table 5** shows that TCQI precedes other learning-based FR-IQA models in terms of prediction accuracy on TID2013 and CSIQ databases. TCQI is slightly worse than the best method on TID2008 and LIVE databases. It is noticed that no IQA approach works very well on all databases except for TCQI. Compared with other learning-based models, although they may work well on one database, fail to predict good result on other databases. For example, [24] and CF-MMF [26] perform well on LIVE and TID2008 databases, respectively. But they cannot provide the best results on TID2013 and CSIQ databases.

Moreover, **Fig. 5** presents the scatter plots of MOS values versus the objective quality scores for TCQI and eleven representative FR-IQA models along with the best fitting logistic functions on TID2013 database. The vertical and horizontal axis denotes the MOS values and objective quality score, respectively. Since this paper proposed a learning-based IQA method by using RF, in order to achieve more stable performance, the training/testing processes are repeated 1000 times, the median result and corresponding training model are retained. Afterwards, we utilize the training model which can produce median result to predict the objective quality scores of all images in TID2013 database. And each blue “+” represents one image in TID2013 database, black curve is a nonlinear logistic fitting curve which obtained by nonlinear fitting according Eq. (21). It is shown in **Fig. 5** that the scatter plots of TCQI are

closer to the black curve compared to other methods. This indicates that TCQI has a higher consistency with subjective perception. For low-quality images, the proposed method can provide the better prediction.

Table 5. Performance comparison of 5 learning-based IQA methods on four databases.

		LLM	[24]	CF-MMF	DOG-SSIMc	TCQI
TID 2013	PLCC	0.9068	0.8939	0.9249	0.9284	0.9448
	SROCC	0.9037	0.8802	0.9143	0.9143	0.9339
	KROCC	0.7209	0.6953	0.7611	0.7542	0.7821
	RMSE	0.5277	0.5527	0.4714	0.4606	0.4030
TID 2008	PLCC	0.8971	0.8881	0.9476	0.9270	0.9428
	SROCC	0.9077	0.8880	0.9422	0.9235	0.9392
	KROCC	0.7368	0.7025	0.8864	0.7721	0.7966
	RMSE	0.5982	0.6121	0.4289	0.5036	0.4426
CSIQ	PLCC	0.9000	0.9680	0.9675	0.9445	0.9766
	SROCC	0.9050	0.9627	0.9668	0.9488	0.9738
	KROCC	0.7238	0.8358	0.8266	0.8128	0.8643
	RMSE	0.1232	0.0650	0.0664	0.0857	0.0532
LIVE	PLCC	0.9578	0.9829	0.9801	0.9515	0.9748
	SROCC	0.9608	0.9755	0.9798	0.9455	0.9721
	KROCC	0.8230	0.8778	0.8574	0.7971	0.8645
	RMSE	7.7678	4.1701	5.4239	4.9045	4.8723

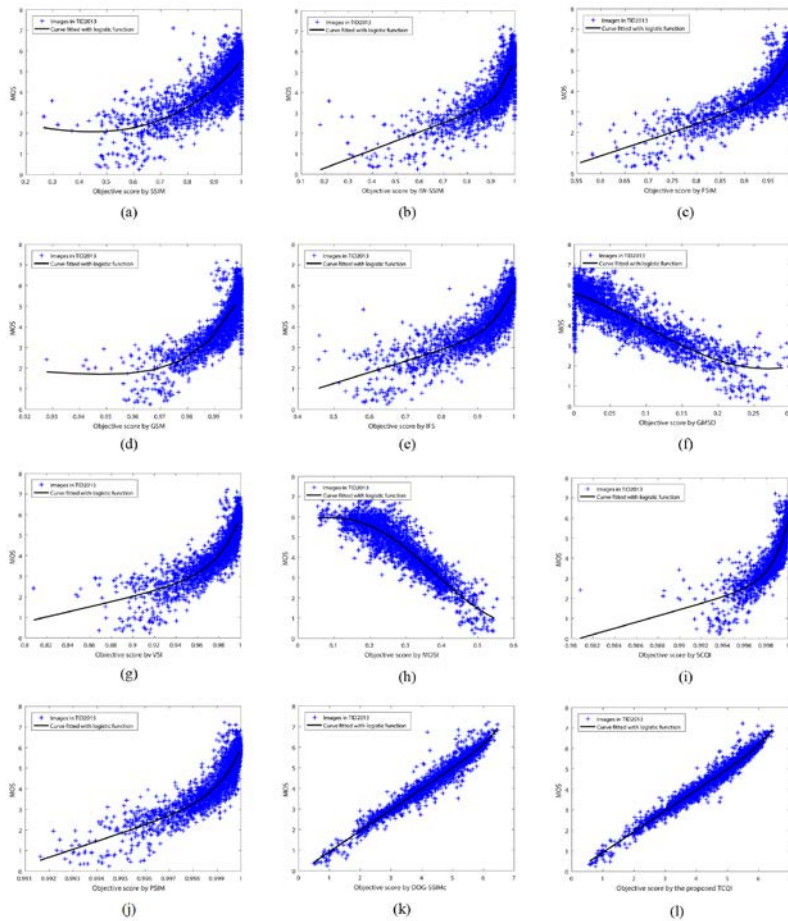


Fig. 5. Scatter plots of objective scores obtained by 12 IQA models versus subjective scores on TID2013 database. (a) SSIM, (b) IW-SSIM, (c) FSIM, (d) GSM, (e) IFS, (f) GMSD, (g) VSI, (h) MDSI, (i) SCQI, (j) PSIM, (k) DOG-SSIMc and (l) TCQI.

In addition, we adopt the F -test to evaluate the statistical significance of TCQI in relative to the competing algorithms, which is based on the residuals between the quality scores obtained by an IQA method after nonlinear regression [38]. As can be seen from Fig. 6, a value of $H=1$ for the left-tailed F -test at a significance level 0.05 denotes that the first algorithm (shown by the row in Fig. 6) is superior in IQA performance to the second algorithm (shown by the column in Fig. 6) with a confidence greater than 95%. On the contrary, $H=0$ indicates that the first algorithm is not significantly better than second algorithm. In other words, $H=0$ represents two algorithm have no significance different in performance. The statistic significant results on TID2013, TID2008, CSIQ and CCID2014 databases are shown in Fig. 6(a-d), respectively. It can be seen that the performances of statistical significance test are consistent with the results shown in Table 4 and Table 5. TCQI is significantly better than other IQA models on TID2013 and CSIQ databases. TCQI, CF-MMF and DOG-SSIMc have no significant difference on TID2008 and LIVE databases. It is noticed that no IQA method performs significantly better than TCQI on all databases.

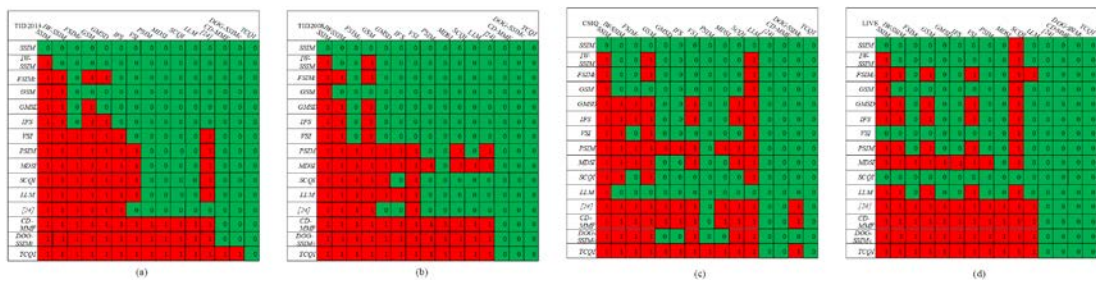


Fig. 6. Illustration of statistical significance tests of competing IQA algorithms on the database of (a) TID2013, (b) TID2008, (c) CSIQ and (d) LIVE. A value ‘1’ denotes (highlighted in red) that the algorithm in the row is significantly better than the algorithm in the column, while a value ‘0’ (highlighted in green) denotes that the first algorithm is not significantly better than the second algorithm.

4.4 Perceptual Features Analysis

In this paper, to examine the respective contribution of each feature relative to the final quality score. We employ the three types of features and randomly concatenate them without any further modifications for the processes of train-test in proposed method TCQI. Then, these features generate the PLCC and SROCC between the predicted quality score and MOS are shown in Fig. 7. The Fig. 7(a) and Fig. 7(b) show the values of PLCC and SROCC, respectively. The dotted lines, solid lines and dash dot lines represent the performance of using only one, two and three types of features, respectively. It can be seen that only two kinds of features are used in regression and the prediction accuracy (PLCC) is beyond 90% on TID2013 and TID2008, 96% on LIVE and CSIQ databases. SROCC can also produce the similar results in Fig. 7(b). Meanwhile, Fig. 7(a) and Fig. 7(b) confirm that the prediction accuracy obtained by using three types of features is significantly better than those obtained by using one or two features.

We also compare the overall performance of different color measurement metrics. Two traditional color measurement metrics (FSIM [9] and VSI [10]) in YIQ and LMN color spaces are compared with proposed color metric. Fig. 8 exhibits that the experimental results in terms of PLCC when using different color spaces. Actually, SROCC can generate the similar results. In Fig. 8, note that the red block and green block represents the predicted results in LMN and

YIQ color spaces, respectively. The blue block denotes the performance of the proposed color metric which is trained by 2-D color features $[\Delta E, \sigma_{\Delta E}]$. It is clear that the prediction accuracy of proposed color metric outperforms the two traditional color metrics on five large-scale databases. The two experimental results indicate that the extracted features in proposed method can also be introduced into existing FR-IQA models.

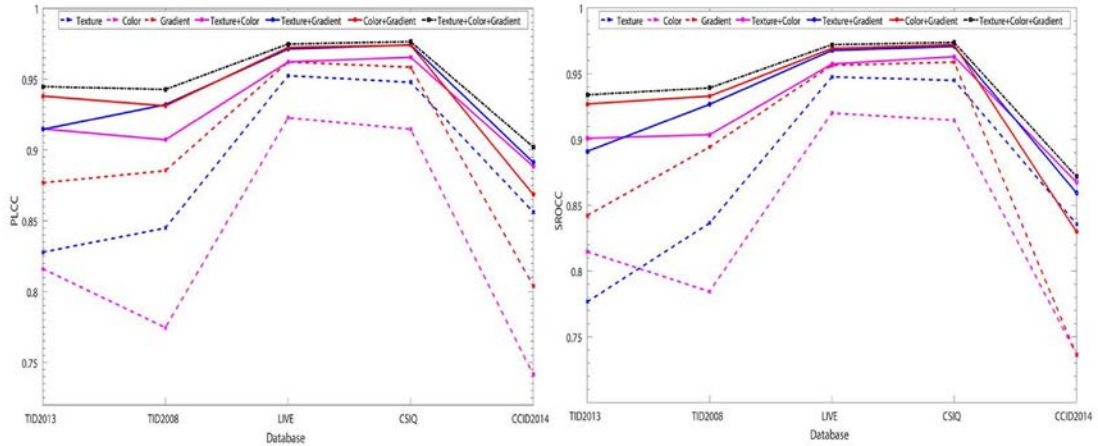


Fig. 7. The respective contribution of each feature relative to the final quality score.

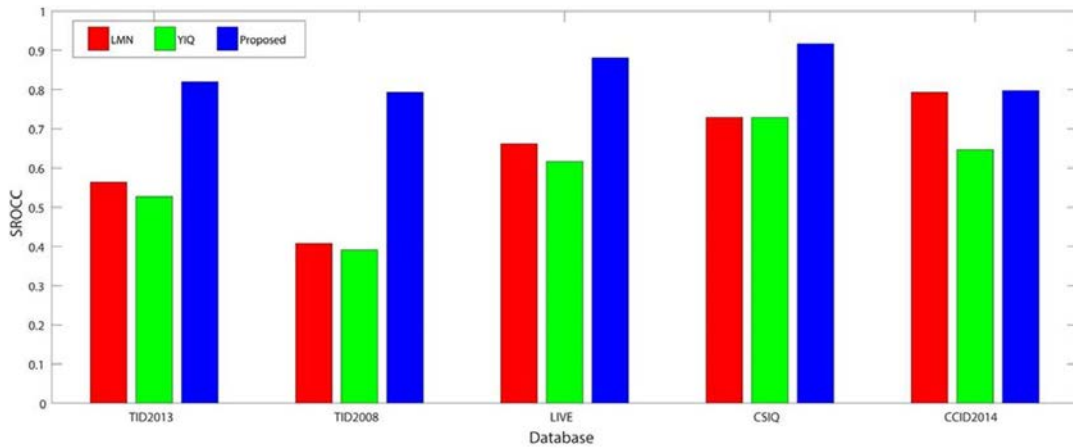


Fig. 8. PLCC values for the different color spaces.

4.5 Regression Approaches for Comparison

Regression tools are critical to the learning-based IQA models. In this paper, performance of RF is compared with the SVR and NN. Table 6 proves that the features extracted by proposed method TCQI can obtain the best results by using RF on all databases. This is why we choose RF as the regression tool in this paper.

Table 6. The performance comparison among different regression tools (RF,SVR and NN)

		PLCC	SROCC	KROCC
TID2013	RF	0.9448	0.9339	0.7821
	SVR	0.8491	0.8530	0.6814
	NN	0.9257	0.9124	0.7496
TID2008	RF	0.9428	0.9392	0.7966
	SVR	0.8458	0.8374	0.6683
	NN	0.9209	0.9181	0.7680
LIVE	RF	0.9748	0.9721	0.8645
	SVR	0.8794	0.8877	0.6980
	NN	0.9604	0.9630	0.8479
CSIQ	RF	0.9766	0.9738	0.8643
	SVR	0.9128	0.9176	0.7547
	NN	0.9724	0.9689	0.8565
CCID2014	RF	0.9021	0.8718	0.6812
	SVR	0.6771	0.6992	0.4934
	NN	0.8601	0.8864	0.6710
Average	RF	0.9470	0.9386	0.7965
	SVR	0.8435	0.8467	0.6716
	NN	0.9280	0.9240	0.7709

4.6 Cross Database Validation

We adopt cross database validation to verify the generality and robustness of proposed method TCQI. One database is used to train regression model and the testing process on another database. The perfect cross validation indicates that the datasets of training and testing should not be overlapping. The same image content exists in TID2013, TID2008, LIVE and CCID2014 databases. Hence, we choose CSIQ as training or testing database and other databases act another role in turn. **Table 7** lists the predicted results in terms of PLCC and the best result obtained by cross database validation is highlighted in red boldface. Actually, the similar performance can also be produced by using SROCC. Compared with five learning-based methods, proposed method acquires the best predicted result on TID2008 and CSIQ databases for all training models. For TID2008 database, the proposed method provides PLCC values of 0.8929 with CSIQ database, which outperforms other learning-based models completely. This is significant, because the size of training dataset is only half of the testing dataset in this case.

Table 7. PLCC values of cross database validation

Training database	TID2008	LIVE	CSIQ	CSIQ
Testing database	CSIQ	CSIQ	TID2008	LIVE
[27]	0.9265	0.9777	0.8608	0.9107
SVDR	0.8831	0.8581	0.7550	0.9086
CD-MMF	0.9127	0.9710	0.8536	0.9109
[24]	0.8863	0.8620	0.7761	0.8852
DOG-SSIMc	0.9154	0.8924	0.8554	0.8934
TCQI	0.9441	0.9545	0.8929	0.9541

4.7 Computational Complexity

The efficiency is another important factor for a good IQA model. To measure the complexity of IQA models, the average running time of all methods are measured for each image in the TID2013 database, which contains 3000 images of size 512×384. Average running time represents the average time consumed to predict the objective quality score for each image in the TID2013 database. The average running time is calculated by using MATLAB functions tic and toc. Experiments are conducted on a computer with 3.60 GHz Intel Core i7 CPU and 8GB RAM. Meanwhile, the MATLAB codes of all other IQA models were obtained from original authors. **Table 8** records the average computation time in seconds. It can be seen in **Table 8** that the proposed TC-QI takes less time than VSI, IW-SSIM, FSIM, SCQI and DOG-FSIMc. Although SSIM, GMSD, MDSI and DOG-SSIMc are faster than TC-QI, these models provide much worse than TCQI. We can conclude that the proposed TCQI obtains the best performance and maintains a moderate computational complexity.

Table 8. The computation time for different methods on TID2013 database

Method	Time(s)	Method	Time(s)
SSIM	0.0338	IW-SSIM	0.2317
VSI	0.1144	FSIM	0.1197
GMSD	0.0163	MDSI	0.0641
SC-QI	0.2617	DOG-SSIMc	0.0521
DOG-FSIMc	0.6590	Proposed TCQI	0.1045

5. Conclusion

In this paper, we firstly reveal that many existing FR-IQA methods can hardly measure the image quality with contrast change. Meanwhile, HVS is sensitive to color degradation and complex texture changes. Based on this, an effective and reliable FR-IQA model for color image is proposed, called TCQI, which combines the color difference in CIE $L^*a^*b^*$ color model and masking texture. Afterwards, we take the gradient feature in spatial domain as the complementary aspects to capture structural information in image. Furthermore, we pool the extracted features with RF procedure toward the objective quality score. Extensive experimental results on five large-scale databases show proposed method has remarkable and robust performances when compared with other 12 state-of-the-art FR-IQA models. Meanwhile, the proposed method correlates well with subjective perception and maintains a moderate computational complexity.

References

- [1] Z. Wang, "Applications of Objective Image Quality Assessment Methods [Applications Corner]," *IEEE Signal Processing Magazine*, vol. 28, no. 6, pp. 137-142, November, 2011. [Article \(CrossRef Link\)](#)
- [2] S. Wang, K. Gu, K. Zeng, Z. Wang, and W. Lin, "Objective Quality Assessment and Perceptual Compression of Screen Content Images," *IEEE Computer Graphics and Applications*, vol. 38, no. 1, pp. 47-58, January, 2018. [Article \(CrossRef Link\)](#)
- [3] K. Gu, D. Tao, J. F. Qiao, and W. Lin, "Learning a No-Reference Quality Assessment Model of Enhanced Images With Big Data," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 29, no. 4, pp. 1301-1313, April, 2018. [Article \(CrossRef Link\)](#)

- [4] K. Gu, J. Zhou, J. F. Qiao, G. Zhai, W. Lin, and A. C. Bovik, "No-Reference Quality Assessment of Screen Content Pictures," *IEEE Transactions on Image Processing*, vol. 26, no. 8, pp. 4005-4018, August, 2017. [Article \(CrossRef Link\)](#)
- [5] W. Zhou, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600-612, April, 2004. [Article \(CrossRef Link\)](#)
- [6] Z. Wang, E. P. Simoncelli, and A. C. Bovik, "Multiscale structural similarity for image quality assessment," in *Proc. of 37th Asilomar Conference on Signals, Systems & Computers*, pp. 1398-1402, November 9-12, 2003. [Article \(CrossRef Link\)](#)
- [7] Z. Wang and Q. Li, "Information Content Weighting for Perceptual Image Quality Assessment," *IEEE Transactions on Image Processing*, vol. 20, no. 5, pp. 1185-1198, May, 2011. [Article \(CrossRef Link\)](#)
- [8] Z. Tang, Y. Zheng, K. Gu, K. Liao, W. Wang, and M. Yu, "Full-Reference Image Quality Assessment by Combining Features in Spatial and Frequency Domains," *IEEE Transactions on Broadcasting*, vol. 65, no. 1, pp. 138-151, March, 2019. [Article \(CrossRef Link\)](#)
- [9] L. Zhang, L. Zhang, X. Mou, and D. Zhang, "FSIM: A Feature Similarity Index for Image Quality Assessment," *IEEE Transactions on Image Processing*, vol. 20, no. 8, pp. 2378-2386, August, 2011. [Article \(CrossRef Link\)](#)
- [10] L. Zhang, Y. Shen, and H. Li, "VSI: A Visual Saliency-Induced Index for Perceptual Image Quality Assessment," *IEEE Transactions on Image Processing*, vol. 23, no. 10, pp. 4270-4281, October, 2014. [Article \(CrossRef Link\)](#)
- [11] W. Xue, L. Zhang, X. Mou, and A. C. Bovik, "Gradient Magnitude Similarity Deviation: A Highly Efficient Perceptual Image Quality Index," *IEEE Transactions on Image Processing*, vol. 23, no. 2, pp. 684-695, February, 2014. [Article \(CrossRef Link\)](#)
- [12] H. Z. Nafchi, A. Shahkolaei, R. Hedjam, and M. Cheriet, "Mean Deviation Similarity Index: Efficient and Reliable Full-Reference Image Quality Evaluator," *IEEE Access*, vol. 4, pp. 5579-5590, October, 2016. [Article \(CrossRef Link\)](#)
- [13] A. Borji and L. Itti, "State-of-the-Art in Visual Attention Modeling," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 1, pp. 185-207, January, 2013. [Article \(CrossRef Link\)](#)
- [14] X. Fei, L. Xiao, Y. Sun, and Z. Wei, "Perceptual image quality assessment based on structural similarity and visual masking," *Signal Processing: Image Communication*, vol. 27, no. 7, pp. 772-783, August, 2012. [Article \(CrossRef Link\)](#)
- [15] W. Lu, T. Xu, Y. Ren, and L. He, "On combining visual perception and color structure based image quality assessment," *Neurocomputing*, vol. 212, pp. 128-134, November, 2016. [Article \(CrossRef Link\)](#)
- [16] D. Lee and K. N. Plataniotis, "Towards a Full-Reference Quality Assessment for Color Images Using Directional Statistics," *IEEE Transactions on Image Processing*, vol. 24, no. 11, pp. 3950-3965, November, 2015. [Article \(CrossRef Link\)](#)
- [17] L. Yang, M. Egawa, M. Akimoto, and M. Miyakawa, "An Imaging Colorimeter for Noncontact Skin Color Measurement," *Optical Review*, vol. 10, no. 6, pp. 554-561, November, 2003. [Article \(CrossRef Link\)](#)
- [18] C. Yang and S. H. Kwok, "Efficient gamut clipping for color image processing using LHS and YIQ," *Optical Engineering*, vol. 42, no. 3, pp. 701-711, March, 2003. [Article \(CrossRef Link\)](#)
- [19] K. Gu, W. Lin, G. Zhai, X. Yang, W. Zhang, and C. W. Chen, "No-Reference Quality Metric of Contrast-Distorted Images Based on Information Maximization," *IEEE Transactions on Cybernetics*, vol. 47, no. 12, pp. 4559-4565, December, 2017. [Article \(CrossRef Link\)](#)
- [20] K. Gu et al., "Saliency-Guided Quality Assessment of Screen Content Images," *IEEE Transactions on Multimedia*, vol. 18, no. 6, pp. 1098-1110, June, 2016. [Article \(CrossRef Link\)](#)
- [21] K. Gu, L. Li, H. Lu, X. Min, and W. Lin, "A Fast Reliable Image Quality Predictor by Fusing Micro- and Macro-Structures," *IEEE Transactions on Industrial Electronics*, vol. 64, no. 5, pp. 3903-3912, May, 2017. [Article \(CrossRef Link\)](#)

- [22] H. Wang, J. Fu, W. Lin, S. Hu, C. C. J. Kuo, and L. Zuo, "Image Quality Assessment Based on Local Linear Information and Distortion-Specific Compensation," *IEEE Transactions on Image Processing*, vol. 26, no. 2, pp. 915-926, February, 2017. [Article \(CrossRef Link\)](#)
- [23] J. Kim, H. Zeng, D. Ghadiyaram, S. Lee, L. Zhang, and A. C. Bovik, "Deep Convolutional Neural Models for Picture-Quality Prediction: Challenges and Solutions to Data-Driven Image Quality Assessment," *IEEE Signal Processing Magazine*, vol. 34, no. 6, pp. 130-141, November, 2017. [Article \(CrossRef Link\)](#)
- [24] Y. Ding, Y. Zhao, and X. Zhao, "Image quality assessment based on multi-feature extraction and synthesis with support vector regression," *Signal Processing: Image Communication*, vol. 54, pp. 81-92, May, 2017. [Article \(CrossRef Link\)](#)
- [25] M. Narwaria and W. Lin, "SVD-Based Quality Metric for Image and Video Using Machine Learning," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 42, no. 2, pp. 347-364, April, 2012. [Article \(CrossRef Link\)](#)
- [26] T. J. Liu, W. Lin, and C. C. J. Kuo, "Image Quality Assessment Using Multi-Method Fusion," *IEEE Transactions on Image Processing*, vol. 22, no. 5, pp. 1793-1807, May, 2013. [Article \(CrossRef Link\)](#)
- [27] S. Wang, C. Deng, W. Lin, G. B. Huang, and B. Zhao, "NMF-Based Image Quality Assessment Using Extreme Learning Machine," *IEEE Transactions on Cybernetics*, vol. 47, no. 1, pp. 232-243, January, 2017. [Article \(CrossRef Link\)](#)
- [28] L. Breiman, "Random Forests," *Machine Learning*, vol. 45, no. 1, pp. 5-32, October, 2001. [Article \(CrossRef Link\)](#)
- [29] S. C. Pei and L. H. Chen, "Image Quality Assessment Using Human Visual DOG Model Fused With Random Forest," *IEEE Transactions on Image Processing*, vol. 24, no. 11, pp. 3282-3292, November, 2015. [Article \(CrossRef Link\)](#)
- [30] S. Bosse, D. Maniry, K. R. Müller, T. Wiegand, and W. Samek, "Deep Neural Networks for No-Reference and Full-Reference Image Quality Assessment," *IEEE Transactions on Image Processing*, vol. 27, no. 1, pp. 206-219, January, 2018. [Article \(CrossRef Link\)](#)
- [31] K. Gu, J. Qiao, X. Min, G. Yue, W. Lin, and D. Thalmann, "Evaluating Quality of Screen Content Images Via Structural Variation Analysis," *IEEE Transactions on Visualization and Computer Graphics*, vol. 24, no. 10, pp. 2689-2701, October, 2018. [Article \(CrossRef Link\)](#)
- [32] H. Haussecker, B. Jähne, and P. Geibler, *Handbook of Computer Vision and Applications with Cdrom*, Morgan Kaufmann Publishers Inc., San Francisco, United States, April, 1999. [Article \(CrossRef Link\)](#)
- [33] N. Ponomarenko, O. Ieremeiev, V. Lukin, K. Egiazarian, L. Jin, J. Astola, B. Vozel, K. Chehdi, M. Carli, F. Battisti and C.-C. Jay Kuo, "Color image database TID2013: peculiarities and preliminary results," in *Proc. of the 4th European Workshop on Visual Information Processing*, pp. 106-111, June 10-12, 2013. [Article \(CrossRef Link\)](#)
- [34] N. Ponomarenko, V. Lukin, A. Zelensky, K. Egiazarian, M. Carli, and F. Battisti, "TID2008 - a database for evaluation of full-reference visual quality assessment metrics," *Advances Modern Radioelectron*, vol. 10, no. 4, pp. 30-45, January, 2009. [Article \(CrossRef Link\)](#)
- [35] E. C. Larson and D. M. Chandler, "Most apparent distortion: full-reference image quality assessment and the role of strategy," *Journal of Electronic Imaging*, vol. 19, no. 1, pp. 1-21, March, 2010. [Article \(CrossRef Link\)](#)
- [36] Z. W. H. R. Sheikh, L. Cormack, and A. C. Bovik., "Live Image Quality Assessment Database Release 2," 2014. available online: <http://live.ece.utexas.edu/research/quality>
- [37] K. Gu, G. Zhai, W. Lin, and M. Liu, "The Analysis of Image Contrast: From Quality Assessment to Automatic Enhancement," *IEEE Transactions on Cybernetics*, vol. 46, no. 1, pp. 284-297, January, 2016. [Article \(CrossRef Link\)](#)
- [38] VQEG, "Final report from the video quality experts group on the validation of objective models of video quality assessment-Phase II," 2013. available online: <https://www.its.bldrdoc.gov/vqeg/projects/frtv-phase-ii/frtv-phase-ii.aspx>

- [39] A. Liu, W. Lin, and M. Narwaria, "Image Quality Assessment Based on Gradient Similarity," *IEEE Transactions on Image Processing*, vol. 21, no. 4, pp. 1500-1512, April, 2012. [Article \(CrossRef Link\)](#)
- [40] H.-w. Chang, Q.-w. Zhang, Q.-g. Wu, and Y. Gan, "Perceptual image quality assessment by independent feature detector," *Neurocomputing*, vol. 151, no. Part 3, pp. 1142-1152, March, 2015. [Article \(CrossRef Link\)](#)
- [41] S. H. Bae and M. Kim, "A Novel Image Quality Assessment With Globally and Locally Consilient Visual Quality Perception," *IEEE Transactions on Image Processing*, vol. 25, no. 5, pp. 2392-2406, May, 2016. [Article \(CrossRef Link\)](#)



Zhisen Tang received a B.E. degree in Printing Engineering from the Hunan University of Technology, Zhuzhou, China in 2016. He is currently pursuing an M.E. degree with the Faculty of Printing, Packaging Engineering and Digital Media Technology, Xi'an University of Technology, Xi'an, China. His research interests include image quality assessment and machine learning.



Yuanlin Zheng received a B.S. degree in Printing Engineering from Zhuzhou Institute of Technology, Zhuzhou, China in 1999, an M.S. degree in Pulp and Papermaking Engineering from Xi'an University of Technology, Xi'an, China in 2002, and a Ph.D. degree in Pulp and Papermaking Engineering from Tianjin University of Science & Technology, Tianjin, China in 2007. He is currently a full associate professor with Printing Engineering at Xi'an University of Technology, Xi'an, China. His research interests include color management, evaluation of quality of color image and color science, and pattern recognition.



Wang Wei received a B.E. degree from the Faculty of Printing, Packaging Engineering and Digital Media Technology, Xi'an University of Technology, Xi'an, China. He is currently pursuing an M.E. degree with Xi'an University of Technology, Xi'an, China. His main research interests include image quality assessment and machine learning.



Kaiyang Liao received a B.S. degree in computer science from the XIDIAN University, Xi'an, China in 2004, an M.S. degree in computer science from the University of Science and Technology LiaoNing, Anshan, China, and a Ph.D. degree in Information and Communication Engineering from the Xi'an Jiaotong University, Xi'an, China in 2013. He is currently a Full lecturer with the Faculty of Printing, Packaging Engineering and Digital Media Technology, Xi'an University of Technology, Xi'an, China. His research interests include data mining and pattern recognition.