Multimedia Tools and Applications
https://doi.org/10.1007/s11042-021-11788-x

Image quality assessment via multiple features

Xichen Yang[1] · **Tianshu Wang**[2] · **Genlin Ji**[1]

Received: 12 October 2020 / Revised: 27 April 2021 / Accepted: 14 December 2021

Abstract
Multimedia devices are indispensable in the information society. And, image quality highly impacts user experience of multimedia equipment. Therefore, measuring image quality accurately has great application value. The existing image quality assessment (IQA) methods have demonstrated the natural sense statistics and image structural information can measure the degradation of image. However, the generalization ability of individual IQA method is limited. In this paper, we propose a novel no-reference IQA method which is based on multiple features. For each image, we first extract natural sense statistic feature, global structural feature and local structural feature, respectively. Second, we train the quality prediction model via different features, and obtain different quality prediction scores by the models. Third, the prediction scores are collected and transformed to feature vectors. Subsequently, the IQA model is trained by support vector regression, and the input variables are the obtained feature vectors and subjective scores. The experimental results on the public databases demonstrate the proposed method can accurately predict the quality of both natural image and screen content image, and the performance is competitive with prevalent methods.

Keywords No reference · Structural information · Natural sense statistic · Image quality assessment · Support vector regression

1 Introduction

Digital images contain a large amount of valid information and have been employed in a variety of daily services [6], such as image retrieval [5, 22], object recognition [2, 3], image steganography [27, 28] and facial emotion recognition [1]. During image acquisition [16, 25], image enhancement [9] and other image processing processes, the quality degradation is inevitable which lead to the loss of effective information in the image [58].

✉ Xichen Yang
xichen_yang@njnu.edu.cn

[1] School of Computer and Electronic Information /School of Artificial Intelligence, Nanjing Normal University, No. 1 Wenyuan Road Qixia District, Nanjing, China

[2] School of Artificial Intelligence and Information Technology, Nanjing University of Chinese Medicine, No. 138 Xianlin Road Qixia District, Nanjing, China

Therefore, evaluating the image quality effectively is of great significance to measure the image usability on the user side. In general, image quality assessment (IQA) method can be divided into subjective IQA method and objective IQA method. For subjective IQA method, image quality is judged via the observation of quality-aware characteristics, which is time-consuming and unstable. With the explosive growth of image in digital services, the subjective IQA is inapplicable, and the objective IQA methods which can automatically measure image quality by computing device are gaining lots of attention [17].

Depending on the amount of information required in quality prediction, objective IQA methods can be divided into full-reference (FR) IQA method, reduced-reference (RR) IQA method and no-reference (NR) IQA method. As the name implies, NR IQA methods need the least information while IQA. Therefore, NR IQA methods are more suitable in practical application. Early NR IQA methods are focus on specific distortion, which should determine the distortion type in advance, and can only apply to restricted distortion types [20]. These drawbacks limit the application of such NR IQA methods.

In practical application, NR IQA methods are usually required to handle multiple and even unknown distortion types [54]. Therefore, a variety of universally purposed NR IQA methods have been proposed. The blind image quality index (BIQI) [35] is a typical NR IQA method. BIQI first estimates the distortions contain in the image, and measures the image quality along each distortion. Finally, BIQI gives out the quality prediction results via probability-weighted summation method. The pre-attention and spatial dependency driven quality assessment (PSQA) predictor [30] employs the pre-attention theory and spatial dependency in IQA. The quality-aware features have significant impact on the accuracy of IQA methods, thus, the researchers pay attention on finding more effective features. The blind/referenceless image spatial quality evaluator (BRISQUE) [33] uses natural scene statistics (NSSs) to measure the losses of naturalness in the image, and then predicts image quality. In [31], the authors try to quantify the degradation of image quality through variations of the structure, naturalness and the perception quality of the distortion image from the pristine image. In addition to demonstrate the effective of NSS features, the experimental results in [31] also indicate the suitable of structural features in IQA. In [53], the proposed method first conducts a model to measure the structural variations in images, and evaluates the image quality by combining two aspects: the distribution of different structural variations and the degree of structural differences. The no reference quality assessment method by incorporating statistical luminance and texture features (NRLT) [8] have demonstrated the suitable of structural information in screen content images (SCIs). Meanwhile, the structural features have been successfully employed in a large amount of image processing algorithms, i.e. the recognition of offline handwritten characters [23], 2D-object recognition [24] and image steganography [29], the structural information closely linked with image usability.

Most of the existing NR IQA methods require training procedure to obtain IQA models, and the support vector regression (SVR) is the most common algorithm. The excellent results demonstrate SVR can effectively establishes the mapping relation between different features and subjective scores for different image type, i.e. stereoscopic image [32], 360° image [19] and deep generative image [21].

The existing IQA methods are mainly focus on single image type. The structural similarity index (SSIM) [57] and the feature similarity index (FSIM) [55] are two typical FR IQA methods that mainly employed in natural image. Besides natural image, a number of FR IQA methods are proposed to apply to screen content image (SCI), i.e. the saliency-guided quality measure of SCIs (SQMS) [11] and the structural variation-based quality index (SVQI) [14]. For NR IQA method, the typical method BRISQUE [33] and NIQE

[34] are focus on natural image. Meanwhile, the big data learning theory has also been employed in SCI quality assessment and one of the typical method is the accelerated screen image quality evaluator (ASIQE) [13]. In [11, 12], the performance comparison results of the existing FR and NR IQA methods on SCIs are illustrated. Despite obtaining excellent results on natural images, the results in [11, 12] demonstrate the existing FR and NR IQA methods are not suitable for SCIs. There are obvious differences between natural images and SCIs. Natural images contain rich and complex distributions of luminance and colour information. SCIs are more complicated when compared with natural images, since SCIs contain both natural scene areas, document areas and graphic areas. In practical application, natural images and SCIs are applied in numerous consumer scenarios, i.e. online game, remote control and cloud computing. Therefore, the IQA method which can apply to both natural image and SCI is more comprehensive and fairer to measure the quality of service on the user side.

The critical step to design IQA method for both natural image and SCI is finding quality-aware features that are sensitive to the degeneration in both the two image types. The main purpose of designing an IQA method is to gain the quality prediction results that are highly consistent with human visual system (HVS). The HVS appears to have both evolved to seek an efficient representation of image information and to match the NSS [44]. And the excellent results obtained by the existing NSS-based NR IQA methods have demonstrated the suitable of NSS features. Meanwhile, the HVS is sensitive to the variation of the structural information in image, and the structural features have been successfully employed in NR IQA. A certain image can be regarded as a three-dimensional space, and the surface can be used to describe the structural information in image. Therefore, the gray-scale fluctuation (GF) of the pixels in different regions can be used for the structural information analysis. The features extracted via GF analysis have been demonstrated effective in predicting the quality of natural image and SCI [49, 50]. Therefore, combining the NSS features and structural features may gain more universal and higher performance on both natural image and SCI.

According to the above discussion, we can draw the following conclusion: (i) NSS features and structural features are effective in IQA; (ii) SVR is suitable for IQA model training; (iii) the IQA method suitable for both natural image and SCI is more applicable. Therefore, we propose a novel NR IQA method via multiple features which can be suitable for both natural image and SCI. The features extracted in the proposed method are NSS features, global structural features and local structural features. The features are combined together to describe the variations of image quality. The procedure of the proposed method can be divided into the following steps. First, the GF analysis method introduced in [49] has employed to calculate the GF features in terms of intensity and direction. Secondly, the NSS features in images are extracted. Thirdly, SVR is employed to conduct a multi-stage learning framework and train the IQA model. The performance of the proposed method is compared on six publicly accessible image databases: LIVE [40], TID2013 [36], LIVE multiply distorted image database (LIVEMD) [18], LIVE in the wild image quality challenge database (LIVEwild) [10], SIQAD [48], and QACS [45]. The experimental results on the six databases have demonstrate the effective of the proposed method on both natural image and SCI in diversely distortion types.

The main contributions of the proposed method can be summarized as blew:

1) The NSS features and structural features are integrated together to gain a comprehensive description of quality variation;

2) The multi-stage training framework via SVR are employed to take advantage of each feature set;
3) The proposed method is suitable for both natural image and SCI, and this advantage ensure the practical value.

The remaining part of this paper is organised as follows. In Sect. 2, the related works are briefly reviewed. The details of the proposed method are described in Sect. 3. The experimental results are illustrated and analysed in Sect. 4. Finally, in Sect. 5, the overall conclusion of the proposed method is provided.

2 Related work

2.1 Image quality assessment via NSS feature

The NSS features have been determined high-performance, and then been employed in a number of IQA methods [41]. For BRISQUE [33], the losses of naturalness in the image are measured via NSS features. In [37], a novel general-purpose NR IQA method which utilising hybrid NSS features is proposed, the method combines the statistical characteristics of mean subtracted and contrast normalised (MSCN) coefficients. The natural image quality evaluator (NIQE) [34] measures the image quality via the distance between the NSS features from a corpus of undistorted images to the NSS features extracted from the distorted images. On the basis of NIQE, the integrated local NIQE (ILNIQE) [56] employs more NSS features, and a Bhattacharyya-like distance is used to measure the image quality. In [42], the IQA method is addressed from two respects, NSS and local sharpness, and associated three feature types. In [47], the NR IQA method is proposed via deep neural networks-based multi-task learning approach, and the proposed network is designed by a multi-task learning manner that consists of NSS features prediction task and the quality score prediction task. The blind IQA method via multiscale natural scene statistical analysis (MNSS) employ two NSS models and achieves better performance than the state-of-the-art quality methods [15]. In [26], the authors propose parametric models to describe the statistical characteristics of chromatic data in natural images, and then quantify visual discomfort caused by chromatic distortions. The experimental results in [37] demonstrate the coefficients in the codebook space keep the NSS characteristics as same as these in the spatial domain, and the distributions of the coefficients can predict the image quality effectively. In [31], the NSS features are extracted via the distributions variations of the locally MSCN coefficients and the products of pairs of the adjacent MSCN coefficients. Motivated by the high performance of the NSS feature, the proposed method employs NSS feature in IQA.

2.2 Image quality assessment via structural feature

In addition to NSS features, in [31] the structural features are captured by the deviations of the image phase congruency and gradients distributions, and the experimental results indicate the structural features can achieve accurate quality prediction results. The effectiveness of the structural features in IQA have been demonstrated in typical FR IQA methods like SSIM [57] and FSIM [55]. Meanwhile, the structural features have gained high performance in various NR IQA methods. In [53], structural variations within images are

classified into four categories: slight deformations, additive impairments, detail losses, and confusing contents, and the quality prediction results. In [8], the NR IQA method incorporates statistical luminance and texture features (NRLT) for measuring the quality of SCIs and gains competitive experimental results. The multiscale local binary pattern (IMLBP) is employed to describe the degeneration of structural characteristics caused by multiply distortions, and then predict the image quality [51]. The directional anisotropic structure measurement (DASM) is introduced in [7]. Based on DASM, the dominant structures in images can be well identified and represented as quality-aware features to measure image quality. The visual pattern degradation based NR IQA model (NRVPD) [46] analyses the visual structure degradations via the orientation similarity based patterns (OSPs), and the quality prediction results are highly consistent with the subjective perception results. In [52], the blind quality assessment method for tone-mapped images is proposed via digging out sufficient information in tone-mapped images, and multiple features are captured to fully represent the colorfulness, naturalness and structure in images. Based on the high applicability of the structural feature in IQA, the proposed method concentrates on combine structural features with NSS features to offer comprehensive description of image quality variations.

2.3 Image quality assessment based on SVR

SVR is widely used in IQA model training. The NR stereoscopic IQA method in [32] firstly extract different features to present the high level semantic perceptual mechanism, image naturalness and the inner binocular interaction information, and then SVR is adopted to train a IQA model based on the extracted features and the subjective scores. The content-aware IQA method for 360° image proposed in [19] employ SVR to train the quality prediction model, and the model is built off-line using a set of training images and corresponding subjective scores. In [21], a multi-stage parallel boosting system is proposed which employ two-stage SVR training procedure. In the first stage, individual features are each used to train separately model by SVR. In the second stage, the prediction scores predicted via the models obtained in the first stage are fed into two SVRs to train the IQA model. The experimental results in [21] demonstrate the suitable of multi-stage SVR training framework. Meanwhile, SVR is widely used in natural IQA [26, 33] and SCI quality assessment [8, 13]. Considering the widely use of SVR in IQA, the proposed method employs SVR in quality prediction model training.

3 Proposed approach

In this section, we detail the proposed NR IQA method. The diversity of distortions and the characteristics of different image types leads to the generalization problem of IQA method. In view of this, the proposed method combines multiple features to describe the degradation caused by the distortions in different views. The framework of the feature extraction of the proposed method is illustrated in Fig. 1. The features extracted in the proposed method are based on the NSS, global structural information and local structural information, respectively. After feature extraction, the features are combined together to gain a comprehensive feature. And then, SVR is employed to conduct a multiple training framework to gain the quality prediction model.

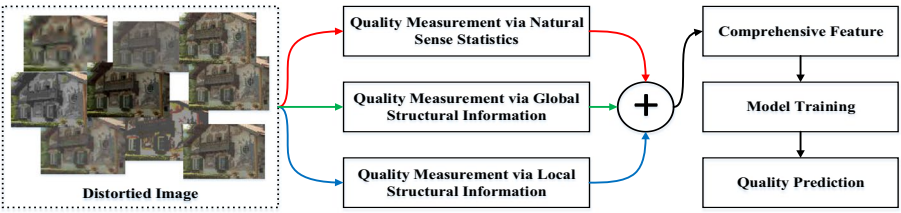

Fig. 1 Theframework of the proposed method

3.1 NSS features

The proposed method extracts NSS features from local image patches to capture the essential low-order statistics of images. Before feature extraction, the image is pre-processed to do local mean removal and divisive normalization. For a certain image I, the pre-processing method is:

$$\widehat{I}(i,j) = \frac{I(i,j) - \mu(i,j)}{\sigma(i,j) + C}, \tag{1}$$

where $i \in \{1,2,\ldots,M\}$, $j \in \{1,2,\ldots,N\}$ are spatial indices, M and N are the dimensions of I, $C = 1$ is a constant that prevents instabilities from occurring when the denominator tends to zero. And,

$$\mu(i,j) = \sum_{k=-K}^{K}\sum_{l=-L}^{L} w_{k,l} I(i+k, j+l), \tag{2}$$

$$\sigma(i,j) = \sqrt{\sum_{k=-K}^{K}\sum_{l=-L}^{L} w_{k,l}[I(i+k, j+l) - \mu(i,j)]^2}, \tag{3}$$

are employed to estimate the local mean and contrast, respectively, where $\mathrm{w} = \{w_{k,l} | k = -K,\ldots,K, l = -L,\ldots,L\}$ is a 2D circularly-sym-metric Gaussian weighting function sampled out to 3 standard deviations. In the proposed method, $K = L = 3$. The transformed luminances $\widehat{I}(i,j)$ calculated by the pre-processing model (1) are refer to mean subtracted contrast normalized (MSCN) coefficients. The MSCN coefficients calculated from natural images have been proved to follow a Gaussian distribution [38]. Different distortions modify the shapes of MSCN coefficients in their own characteristic way. Therefore, the modification degree of MSCN coefficients distribution can measure the distortion severity. Prior researches [57, 58] have demonstrated that the generalized Gaussian distribution (GGD) can effectively captures behaviour of the MSCN coefficients of different distortion images. The GGD with zero mean can be given by:

$$f(x;\alpha,\sigma^2) = \frac{\alpha}{2\beta\Gamma(1/\alpha)} exp\left(-\left(\frac{|x|}{\beta}\right)^{\alpha}\right), \tag{4}$$

where

$$\beta = \sigma \sqrt{\frac{\Gamma(1/\alpha)}{\Gamma(3/\alpha)}}, \tag{5}$$

and $\Gamma(\cdot)$ is the gamma function:

$$\Gamma(a) = \int_0^{\infty} t^{a-1} e^{-t} dta > 0. \tag{6}$$

The parameters of the $GGD(\alpha, \sigma^2)$, can be reliably estimated via the moment-matching based approach in [39]. In all of the parameters, the shape parameter α controls the 'shape' of the distribution, and σ^2 controls the variance. For each image, the parameters α and σ^2 are estimated from a GGD fit of the MSCN coefficients, and employed to conduct quality-aware feature.

The signs of adjacent MSCN coefficients also exhibit a regular structure, and can be disturbed by image distortions. In view of this, the empirical distributions of pairwise products of adjacent MSCN coefficients are analysed along four orientations: horizontal (H), vertical (V), main-diagonal (D1) and secondary-diagonal (D2). The calculation methods are given by:

$$H(i,j) = \widehat{I}(i,j)\widehat{I}(i,j+1), \tag{7}$$

$$V(i,j) = \widehat{I}(i,j)\widehat{I}(i+1,j), \tag{8}$$

$$D1(i,j) = \widehat{I}(i,j)\widehat{I}(i+1,j+1), \tag{9}$$

$$D2(i,j) = \widehat{I}(i,j)\widehat{I}(i+1,j-1), \tag{10}$$

where $i \in \{1,2 \ldots M\}$, $j \in \{1,2 \ldots N\}$. The asymmetric generalized Gaussian distribution (AGGD) with zero mode is employed to captures behaviour of the adjacent MSCN coefficients in different orientations. The AGGD with zero mode is given by:

$$f\left(x;\nu,\sigma_l^2,\sigma_r^2\right) = \begin{cases} \frac{\nu}{(\beta_l+\beta_r)\Gamma\left(\frac{1}{\nu}\right)} exp\left(-\left(\frac{-x}{\beta_l}\right)^{\nu}\right) x < 0 \\ \frac{\nu}{(\beta_l+\beta_r)\Gamma\left(\frac{1}{\nu}\right)} exp\left(-\left(\frac{x}{\beta_r}\right)^{\nu}\right) x \geq 0 \end{cases}, \tag{11}$$

where

$$\beta_l = \sigma_l \sqrt{\frac{\Gamma(1/\nu)}{\Gamma(3/\nu)}}, \tag{12}$$

$$\beta_r = \sigma_r \sqrt{\frac{\Gamma(1/\nu)}{\Gamma(3/\nu)}}. \tag{13}$$

In the formula, ν describe the shape of adjacent MSCN coefficients, and σ_l^2 and σ_r^2 are scale parameters that control the spread on each side of the mode. The parameters ν, σ_l^2 and

σ_r^2 are employed to as quality-aware features. Besides the three parameters, the mean of the distribution is also calculated and defined as η, the calculation method is

$$\eta = (\beta_r - \beta_l)\frac{\Gamma(2/\nu)}{\Gamma(1/\nu)}. \tag{14}$$

The four parameters of AGGD are extracted along the four orientations, respectively. And then, the 16 parameters are obtained to describe the distribution of adjacent MSCN coefficients. Together with the parameters of the $GGD(\alpha, \sigma^2)$, the 18 parameters are employed to conduct NSS feature, and the NSS feature is denoted as *Nfeat*.

3.2 Structural features

The distortions always lead to the variation of image texture information, i.e. smoothing, sparseness, and regularity. The texture variation can be measured effectively via the changes of structural information in local and global regions. Therefore, the degree of distortions can be measured by structural information analysis. In this paper, the GF analysis method in [49, 50] is employed to extract the structural features in local and global regions.

The texture characteristics can be described effectively by the relationship between the gray-scale of adjacent pixels. Therefore, the intensity of GF can be used to measure the degree of texture variation in certain region. As illustrated in Fig. 2a, the GF intensity of the central pixel $I(i,j)$ in its adjacent area is analysed in four detection directions: 0°, 45°, 90° and 135°. The pixels in the four detection directions are given by: $\{I(i,j-1), I(i,j), I(i,j+1)\}$, $\{I(i+1,j-1), I(i,j), I(i-1,j+1)\}$, $\{I(i-1,j), I(i,j), I(i+1,j)\}$ and $\{I(i-1,j-1), I(i,j), I(i+1,j+1)\}$, respectively. The sample of GF intensity calculation in a certain detection direction is illustrated in Fig. 2b. In the coordinate axes, the origin o is the central pixel, and o_1 and o_2 are the neighbours of the central pixel in current detection direction. The coordinates of o_1 and o_2 are $(-1, d_1)$ and $(1, d_2)$, where d_1 and d_2 are given: $d_1 = g(o) - g(o_1)$, $d_2 = g(o) - g(o_2)$. And, $g(\cdot)$ is the gray-scale values of the certain pixel. The θ_1 and θ_2 are the angles between the coordinate and $\overrightarrow{oo_1}$ and $\overrightarrow{oo_2}$. It is obvious that, the GF is proportional to the magnitudes of θ_1 and θ_2, and θ_1 and θ_2 are both range from 0° to 90°. The two angles θ_1 and θ_2 are combined to measure the GF intensity and defined as $\theta = \theta_1 + \theta_2$. And $\cos\theta$ is in the range $[0^\circ, 180^\circ)$. Considering that $\cos\theta$ monotonically decrease in the range $[0^\circ, 180^\circ)$. Therefore, $\cos(\theta + \pi)$ is employed to measure the intensity which monotonically increases in the range $[0^\circ, 180^\circ)$. As illustrated in Fig. 2c, there are four situations on the basis of the positive and negative statuses of d_1 and d_2. It is obvious that, the GF is relatively higher when the positive and negative statuses of d_1 and d_2 are different. The GF intensity of $I(i,j)$ is denoted as $Gi(i,j)$, and its calculation method is given by:

$$Gi(i,j) = \begin{cases} \cos(\theta(i,j) + \pi), d_1 \times d_2 \geq 0, \\ \cos(\theta(i,j) + \pi) + C, d_1 \times d_2 < 0, \\ \theta(i,j) = \theta_1(i,j) + \theta_2(i,j), \end{cases} \tag{15}$$

where C is constant. Considering that $\cos(\theta(i,j) + \pi)$ is in the range $[-1, 1)$, the value of C is assigned 1. On the basis of formula (15), $Gi(i,j)$ is in the range $[-1, 2)$. For the certain image I, formula (15) is employed to calculate the GF intensity of each pixel and obtain the GF map (GFM) in certain detect direction. The values of pixels in GFM are proportional to the GF intensity between the pixels and their neighbours in certain detection direction, and

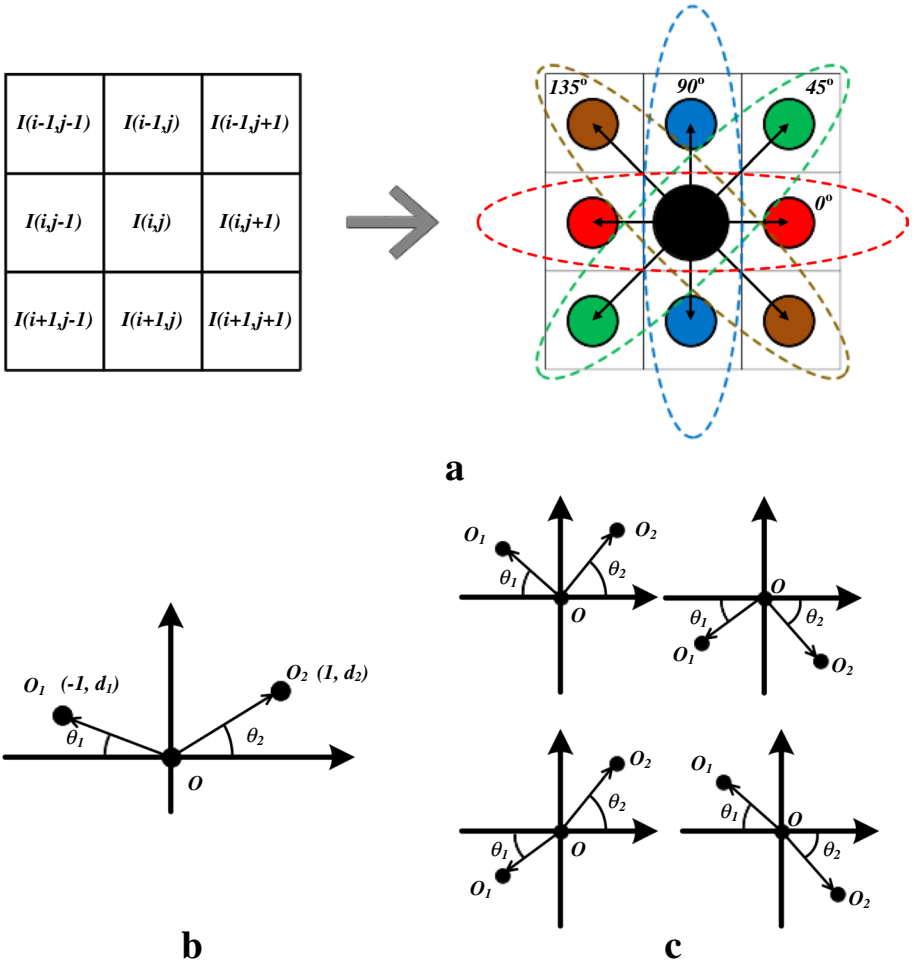

Fig. 2 Gray-scale fluctuation intensity: (**a**) Detection direction, (**b**) Sample in a certain detection direction, (**c**) Different situations of gray-scale fluctuation

the GFM in the detection direction x is defined as Gm_x, where $x \in \{1,2,3,4\}$ corresponds to the detection directions $\{0^\circ, 45^\circ, 90^\circ, 135^\circ\}$.

On the other hand, image distortions also change the direction information in the image. the direction information can be changed by image distortion. Therefore, the proposed method also employs the direction information of GF for distortion assessing. For $I(i,j)$, the intensities of GF in the four detection directions are calculated, and then the direction that correspond to the maximum GF intensity is selected as the direction feature and defined as $x_{max}(i,j)$. The GF direction map (GFD) of I is denoted as Gd. The values in Gd can be obtained after per-pixel calculation via $Gd(i,j) = x_{max}(i,j)$, where $i \in \{1,2,\ldots,M\}$, $j \in \{1,2,\ldots,N\}$, and M and N are the dimensions of I.

After the analysis of GF intensity and direction of image I, we can obtain four GFMs and one GFD. The five feature maps can effectively describe the distribution of structural information in I, and can be employed to extract structure-sensitive features [49].

The proposed method extracts 11 structural features, and the features are denoted as $\{f_1,f_2,f_3,f_4,f_5,f_6,f_7,f_8,f_9,f_{10},f_{11}\}$. Besides GFMs and GFD, the proposed method analyses the mean value, median value and standard deviation value of GF intensity in local regions in GFMs and gain the GF mean map, GF Median map and GF STD map. Additionally, the standard deviation values of local regions in GF mean map, GF Median map and GF STD map are calculated to describe the variations in such regions and gain the three feature maps. The 11 features are histogram-based and describe the distributions of the values in the 11 structural feature maps. The histogram-based feature of a feature map whose size is $M \times N$ is denoted as $Shist = \{h_1, h_2, \dots, h_k\}$, and its calculation method is given by:

$$h_k = \frac{1}{M \times N} \sum_{i=1}^{M} \sum_{j=1}^{N} \Omega(Fm(i,j), H(k)),$$
$$\Omega(a,b) = \begin{cases} 1, a \in b \\ 0, a \notin b \end{cases}, \tag{16}$$

where k is the bin index, $H(k)$ is the interval of each bin, Fm is the feature map, and h_k is the existence probability of certain value in Fm. For f_5, the range of k is $[1,4]$. And k of the other features is in the range of $[1,10]$.

The details of the 11 features are listed in Table 1. For image I, f_1-f_4 can describe the global distribution of the GF intensities in I, and f_5 can describe the global distribution of the GF direction in I. The GF mean map, GF Median map and GF STD map can express the GF intensity in local regions of I, thus, f_6-f_8 can well describe local structural features in I. And, f_9-f_{11} can express the variations of local structural information. The 11 features are combined to form the structural feature $Sfeat = \{f_1,f_2, \dots ,f_{11}\}$, which contain 104 parameters. Considering that, the GFM and GFD can describe both the intensity and direction of image structural information effectively. And, the distributions of the pixel values in both GFM and GFD are sensitive to image distortion. Therefore, *Sfeat* can describe both global structural information and local structural information in I, and is suitable for measuring the distortion contained in image.

3.3 Model training and quality prediction

The proposed method employs a two-stage training procedure to train the quality prediction model. The flow chart of model training is illustrated in Fig. 3. Firstly, *Nfeats* and *Sfeats* are extracted from the images in the first training set. The extracted features are divided into four sets, that are *Nfeat*, $\{f_1,f_2,f_3,f_4,f_5\}$, $\{f_6,f_7,f_8\}$ and $\{f_9,f_{10},f_{11}\}$. Secondly, the implicit relationship between each feature set and the corresponding subjective scores is established by SVR. The models belong to different feature sets are denoted as $Model_1$, $Model_2$, $Model_3$ and $Model_4$. The LibSVM package [4] is employed to implement

Table 1 Details of structural features

Feature ID	Feature description
f_1-f_4	The histogram-based features that describe the distribution of the GF intensities in the GFMs of four detection directions.
f_5	The histogram-based feature that describes the distribution of the GF direction in the GFD.
f_6-f_8	The distributions of values in GF Mean map, GF Median map and GF STD map.
f_9-f_{11}	The distributions of standard deviation analysis results of local regions in GF Mean map, GF Median map and GF STD map.

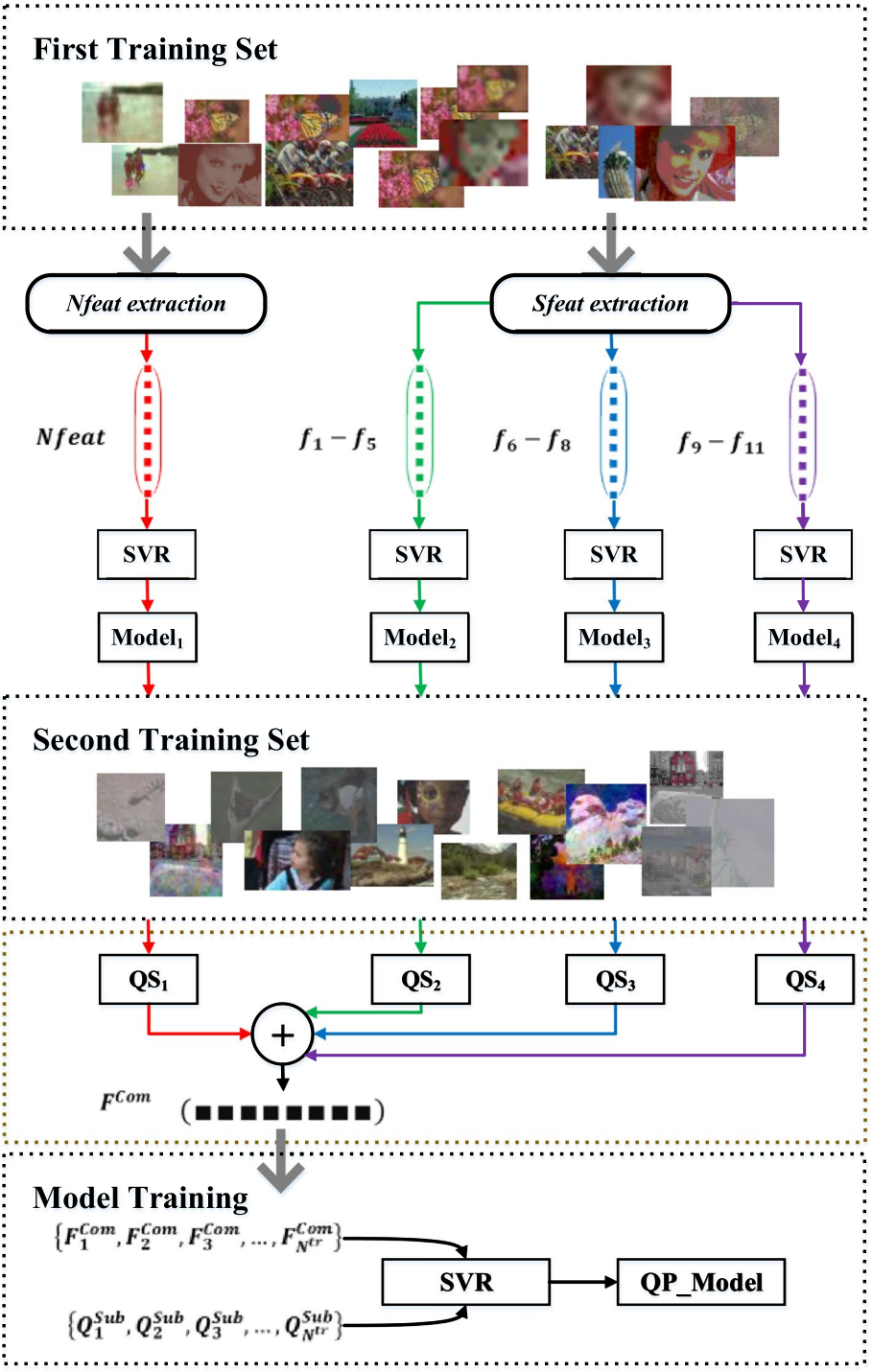

Fig. 3 Flow chart of model training

the model training. Thirdly, the quality scores of the images in the second training set are predicted via $Model_1$, $Model_2$, $Model_3$ and $Model_4$, respectively. Defining a certain image in the second training image set as I_i^{tr}, where $i \in \{1,2,\ldots,N^{tr}\}$ and N^{tr} is the number of images in the training set. For I_i^{tr}, four quality prediction scores corresponding to $Model_1$, $Model_2$, $Model_3$ and $Model_4$ can be obtained, that are QS_{1i}, QS_{2i}, QS_{3i} and QS_{4i}. And then, QS_{1i}, QS_{2i}, QS_{3i} and QS_{4i} are combined together to form the multiple feature of I_i^{tr} which is denoted as F_i^{Com}. Fourthly, the mapping relationship between the multiple features:$\left\{F_1^{Com}, F_2^{Com}, F_3^{Com}, \ldots, F_{N^{tr}}^{Com}\right\}$ and the subjective scores: $\left\{Q_1^{Sub}, Q_2^{Sub}, Q_3^{Sub}, \ldots, Q_{N^{tr}}^{Sub}\right\}$ of the images in the second training set is established via SVR to gain the quality prediction model *QP_Model*.

For a certain distortion image I^{test}, its *Nfeat* and *Sfeat* are firstly extracted. And then, its quality prediction scores which are corresponding to $Model_1$, $Model_2$, $Model_3$ and $Model_4$ can be obtained and denoted as QS_1^{test}, QS_2^{test}, QS_3^{test} and QS_4^{test}. After that, QS_1^{test}, QS_2^{test}, QS_3^{test} and QS_4^{test} are combined to form the comprehensive feature F_{test}^{Com}. Finally, relying on *QP_Model* and F_{test}^{Com}, the quality of I^{test} can be predicted automatically.

4 Experimental results and analysis

4.1 Databases and metrics

The performance comparisons of the proposed method are conducted on six publicly accessible image databases, including: LIVE [40], TID2013 [36], LIVEMD [18], LIVEwild [10], SIQAD [48], and QACS [45]. The images in LIVE, TID2013, LIVEMD and LIVEwild are natural image. On the other hand, the SIQAD and QACS are focus on SCI. The distortions contained in LIVEMD are two multiple distortion scenarios. Meanwhile, the images in LIVEwild are collected from highly diverse mobile devices and contain multiple interacting authentic distortions of different levels. In summary, the image databases contain both natural image and SCI with individual distortion types and multiple distortions. Therefore, the performance comparisons can demonstrate the applicability of the proposed image on both image types and distortion types.

Five performance metrics are employed to measure the performance of the proposed method quantificationally, including Spearman rank-ordered correlation coefficient (SROCC), Pearson linear correlation coefficient (LCC), Kendall rank-order correlation coefficient (KROCC), mean absolute error (MAE), and root-mean-square error (RMSE). The values of SROCC and KROCC can measure the monotonicity of the prediction results. The values of LCC, MAE and RMSE are employed to measure the accuracy of the prediction results. Before calculating the performance metrics, a logistic function [43] is employed for non-linear mapping to remove the non-linearity caused by the subjective quality assessment, and to ensure the comparisons are done in a common analysis space. The logistic function is given by:

$$Q(q) = \beta_1 \left(\frac{1}{2} - \frac{1}{1 + e^{\beta_2 (q - \beta_3)}} \right) + \beta_4 x + \beta_5, \tag{17}$$

where q is the quality prediction score, Q is the corresponding mapped score, and β_1, β_2, β_3, β_4 and β_5 are the parameters of the logistic function.

Before model training, we random select certain proportion of the images in each database to form the first training set, the second training set and the test set, respectively. And,

we ensure no cross-content between the training set and the test set. The random selection can avoid the influence of the training set selection on the performance of the proposed method. The train-test procedure repeats 100 times on each image database, and then the median performance across these iterations are selected as the final IQA result of the proposed method.

4.2 Effect of extracted features

The proposed method divides the extracted features into four feature sets, that are *Nfeat*, $\{f_1,f_2,f_3,f_4,f_5\}$, $\{f_6,f_7,f_8\}$ and $\{f_9,f_{10},f_{11}\}$. We conduct comparison experiments to measure the effect of individual feature set.

The experimental results of different feature sets are illustrated in Fig. 4. In Fig. 4, the first FS, second FS, third FS and fourth FS are corresponding to *Nfeat*, $\{f_1,f_2,f_3,f_4,f_5\}$, $\{f_6,f_7,f_8\}$ and $\{f_9,f_{10},f_{11}\}$, respectively. The results in Fig. 4 can offer us the following information. First, the SROCC and LCC values of the four features sets are both higher or close to 0.9 on LIVE. Second, for all the feature sets, the SROCC values are around 0.7, and the LCC values are close to 0.8 on SIQAD. Third, the KROCC values of all the feature sets are around 0.7 on LIVE, and are higher than 0.5 on SIQAD.

Based on the above discussions, each individual feature set can gain competitive performance on both natural image and SCI. Combining all the four features can offer more

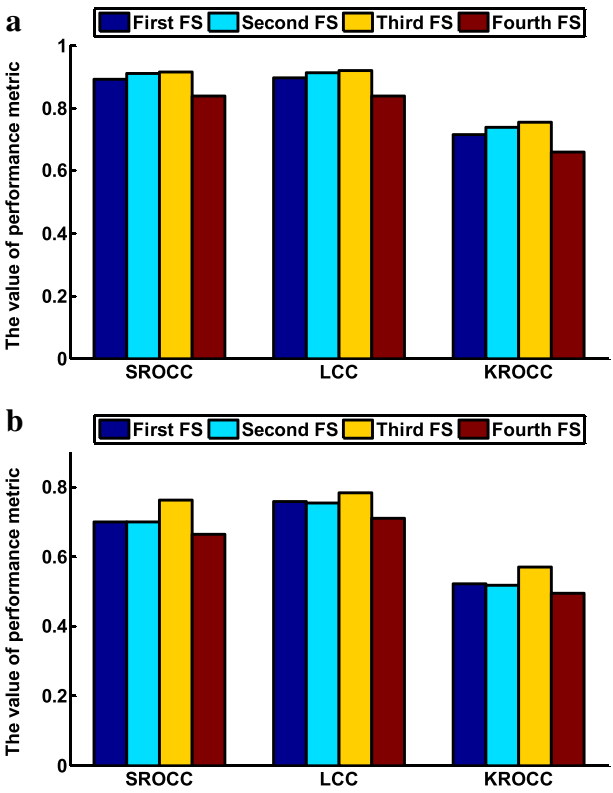

Fig. 4 Performance of different feature sets: (**a**) Performance metrics on LIVE, (**b**) Performance metrics on SIQAD

comprehensive description of image quality. Therefore, the proposed method employs all the feature sets in IQA.

4.3 Effect of different scales

The features extracted via multiple scales can offer more comprehensive description for the characteristics of image quality. The selected scale and the number of scales employed in IQA inevitably affect the performance of the proposed method. Figures 5 and 6 illustrate the performance of different scales and different scale number on LIVE database. In Figs. 5 and 6, the ordinates are the values of performance metrics, and the abscissas identify different distortion types. In Fig. 5, the first scale corresponds to the original image, the second scale corresponds to the original image down sampling via factor 2, and the third scale corresponds to the original image down sampling via factor 4.

The performance comparison results in Figs. 5 and 6 offer us the following information. First, the first scale can obtain the highest performance while the values of SROCC and LCC on all the distortion types are both higher than 0.9. Second, the performance of the second scale is close to the first scale and also higher than 0.9. Third, the overall performance on LIVE are relatively higher while the scale number is 2. Meanwhile, the performances over different distortion types are close when the scale numbers are 1 and 2.

The discussions above have demonstrated that the selected scale and the scale number can influence the accuracy of the proposed method. In the following performance analysis, the proposed method employs the first and second scales for feature extraction.

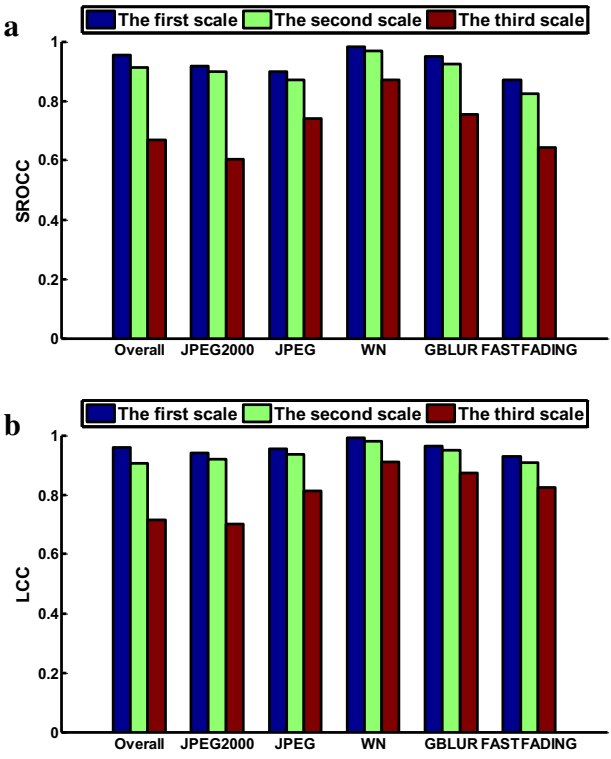

Fig. 5 Performance of different scales: (**a**) SROCC values on LIVE, (**b**) LCC values on LIVE

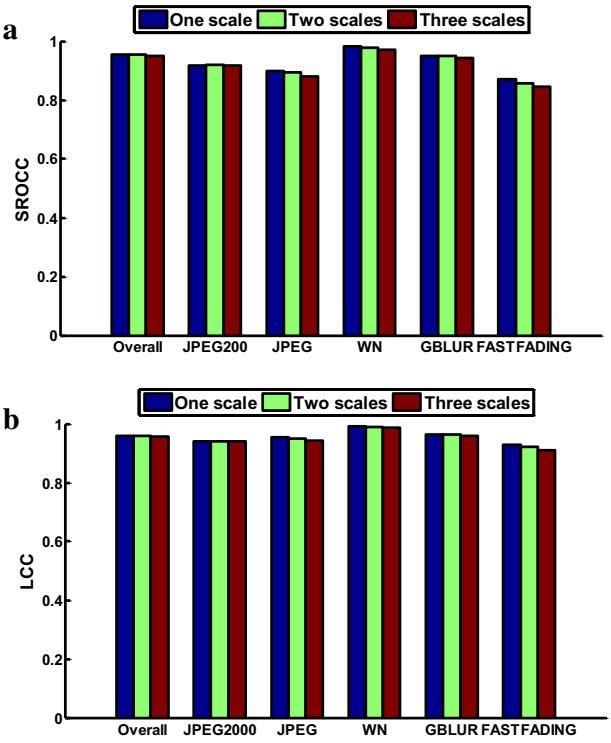

Fig. 6 Performance of multiple scale: (**a**) SROCC values on LIVE, (**b**) LCC values on LIVE

4.4 Sensitivity to training set variation

In real-world application scenario, the available training images are limited. Therefore, the robustness to the scale of training images is important for IQA method. Figures 7 and 8 illustrate the performance of the proposed method of different training scales on LIVE database and SIQAD database. The ordinates are the values of performance metrics, and the abscissas identify distortion types. The proportions of images selected for establishing the first training set and the second training set are 15%, 25%, 35% and 45%. The test set occupy 70%, 50%, 30% and 10%, respectively. For each training proportion, the train-test procedure is repeated 100 times and the median result is calculated as the final performance.

The results in Figs. 7 and 8 lead us to the following conclusions. First, the larger training proportion obtain higher performance on the overall database and each distortion type. Second, on LIVE and SIQAD, even when the training set occupy 15%, the SROCC and LCC values of the overall results are no more than 0.1 lower than the highest training proportion. Third, the performances fluctuate slightly while the training proportion changes on single distortion types.

The training images in practical applications are limited. Therefore, the generalisation capability is important for IQA method. The results of cross database validation on LIVE, TID2013 and SIQAD are illustrated in Table 2. Four distortion types: JPEG2000 compression (JPEG2000), JPEG compression (JPEG), white noise (WN) and Gaussian blur (GBLUR) are selected in the comparison experiments, which are the distortion types

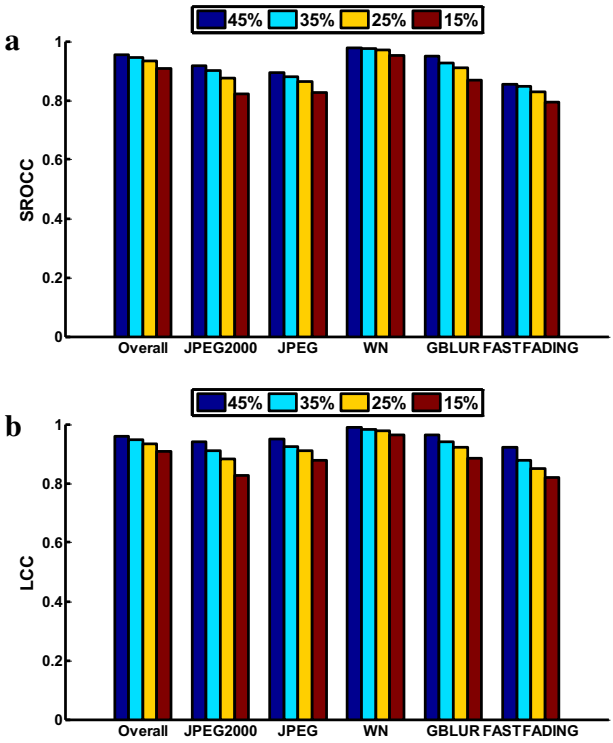

Fig. 7 Performance of different training set proportions on LIVE: (**a**) SROCC values, (**b**) LCC values

shared in the three databases. For the image database selected for model training, we randomly select half reference images and their distorted versions to establish the first training set, and the rest images are selected to establish the second training set. It is notable that the proposed method gains higher performance when the training images and test images are both natural images, the SROCC and LCC values on overall database and individual distortion types are all mostly higher than 0.85. And, there is a bit performance degradation when the image types in the training set and test set are different. However, all the results illustrated in Table 2 demonstrate the competitiveness of the proposed method when the type of training image varies.

The experimental results demonstrate the performance of proposed method can be improved by enlarging the training set. The proposed method can keep stable when the scale of training set changes for both natural image and SCI. And, the proposed method can keep competitive while the training set proportions are relatively small. Meanwhile, the proposed method gains acceptable results when the types of training images and test images are different.

4.5 Performance comparison

The performance of the proposed method is compared with eight IQA methods, which are four FR IQA methods: SSIM [57], FSIM [55], SQMS [11], and SVQI [14], and four NR IQA methods: NRLT [8], NIQE [34], ILNIQE [56], and ASIQE [13]. The eight methods include both FR IQA and NR IQA. SSIM, FSIM, NIQE and ILNIQE are mainly employed

Table 2 Prediction performance on cross-database validation

Training Database	SROCC	Overall	JPEG2000	JPEG	WN	GBLUR	LCC	Overall	JPEG2000	JPEG	WN	GBLUR
TID2013	LIVE	0.8511	0.8924	0.8781	0.8799	0.8773	LIVE	0.8506	0.8974	0.9159	0.8923	0.8757
SIQAD		0.7486	0.7913	0.7197	0.9288	0.8597		0.7601	0.7997	0.7667	0.9308	0.8691
LIVE	TID2013	0.8597	0.9140	0.8820	0.8773	0.8486	TID2013	0.8835	0.9280	0.9015	0.8763	0.8460
SIQAD		0.7622	0.8280	0.7531	0.7706	0.7814		0.7889	0.8751	0.8109	0.7849	0.8021
LIVE	SIQAD	0.7061	0.6359	0.6169	0.8772	0.7725	SIQAD	0.7333	0.6497	0.6159	0.8856	0.7340
TID2013		0.6732	0.5088	0.6421	0.8735	0.6353		0.7098	0.5490	0.6625	0.8908	0.6467

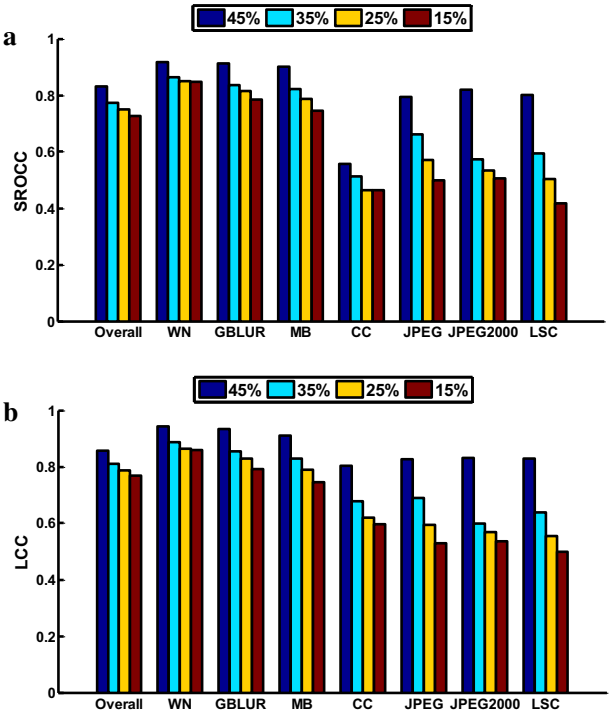

Fig. 8 Performance of different training set proportions on SIQAD: (**a**) SROCC values, (**b**) LCC values

in natural IQA, and SVQI, SQMS, ASIQE and NRLT are for SCI. Meanwhile, the selected methods are either for natural image or SCI. Therefore, the performance comparisons with the eight methods can demonstrate the competitive of the proposed method when the image type or distortion type varies. For each method, 80% images are randomly selected for training and the rest images for performance verification. However, the proposed method requires two training sets and one test set on each database, and the proportion of each part are 45%, 45% and 10%, respectively. And, the train-test procedures are repeated 100 times, and the median value across these iterations are calculated as the final result of the certain method. In this way, the interference caused by training set selection is minimised.

The performances of the selected methods on each database are illustrated in Table 3. And, the average performance of each method across different databases is illustrated in Table 4. Considering that, there are no pristine images in LIVEwild database, thus, LIVEwild database gets sidestepped when calculating the average performance of each method. The average results in Table 4 are calculated in two cases: (i) directly averaged across all databases; (ii) averaged due to the scale of each database. In Tables 3 and 4, the best two results of each metric are highlighted in bold. The results illustrated in Tables 3 and 4 can lead to following information. First, when compared with FR IQA methods, the proposed method gains the highest performance on LIVE, LiveMD1 and Live MD2, and is second only to FSIM on TID2013. The SROCC and LCC values of the proposed method are all exceed 0.84 on QACS and SIQAD, which are competitive with the best two FR IQA methods. Second, the proposed method gains higher performance on both natural image databases and SCI databases when compared with NIQE and NRLT which are focus on NSS features and structural features, respectively. The results demonstrate the effective of comprehensive feature. Meanwhile, the proposed method gains better performance when

Table 3 Performance comparisons on different databases

		FR IQA methods					NR IQA methods				
	Metric	SSIM	FSIM	SVQI	SQMS	PROPOSED	NIQE	ILNIQE	ASIQE	NRLT	PROPOSED
LIVE	SROCC	0.8510	**0.9226**	0.9016	0.9017	**0.9474**	0.8360	0.8472	0.6837	**0.8961**	**0.9474**
	LCC	0.8772	**0.9234**	0.9026	0.9027	**0.9491**	0.8443	0.8549	0.6958	**0.9002**	**0.9491**
	KROCC	0.6845	**0.7534**	0.7209	0.7216	**0.8064**	0.6317	0.641	0.4891	**0.7188**	**0.8064**
	RMSE	7.7312	**6.181**	6.9317	6.9288	**7.2695**	8.6285	8.3536	11.5666	**6.9802**	**7.2695**
	MAE	5.8632	**4.6636**	5.3292	5.2889	**5.3954**	6.5975	6.6028	9.2038	**5.2436**	**5.3954**
TID2013	SROCC	0.6274	**0.8022**	0.6065	0.6789	**0.7194**	0.3119	0.4938	0.3435	**0.5711**	**0.7194**
	LCC	0.6861	**0.8589**	0.7335	0.7619	**0.7800**	0.3991	0.5884	0.5601	**0.653**	**0.7800**
	KROCC	0.4564	**0.6292**	0.4423	0.5024	**0.5395**	0.2119	0.3491	0.268	**0.4132**	**0.5395**
	RMSE	0.9019	**0.6349**	0.8425	0.8029	**0.7789**	1.1367	1.0024	1.027	**0.9325**	**0.7789**
	MAE	0.7263	**0.4825**	0.6975	0.6525	**0.5970**	0.9416	0.8112	0.863	**0.7554**	**0.5970**
LIVEwild	SROCC	-	-	-	-	0.6543	0.4509	0.4397	0.3156	**0.5941**	**0.6543**
	LCC	-	-	-	-	0.7040	0.4989	0.5047	0.3587	**0.6184**	**0.7040**
	KROCC	-	-	-	-	0.4743	0.3079	0.2987	0.2173	**0.4245**	**0.4743**
	RMSE	-	-	-	-	14.4721	17.5745	17.5068	18.9286	**15.8966**	**14.4721**
	MAE	-	-	-	-	11.2043	14.0128	14	15.186	**12.3184**	**11.2043**
LiveMD1	SROCC	0.7443	**0.8546**	0.7892	0.7918	**0.9343**	0.8707	0.8911	0.7221	**0.8936**	**0.9343**
	LCC	0.8003	0.8192	0.8519	**0.853**	**0.9528**	**0.9099**	0.9046	0.8424	0.8962	**0.9528**
	KROCC	0.5512	**0.6606**	0.5878	0.5901	**0.7815**	0.6891	0.7018	0.5418	**0.7172**	**0.7815**
	RMSE	11.4894	10.9883	10.0339	**10.0011**	**5.6628**	**7.9494**	8.1694	10.3258	8.1501	**5.6628**
	MAE	9.1671	9.227	8.1402	**8.128**	**4.5506**	**6.2393**	6.7569	8.1541	6.244	**4.5506**
LiveMD2	SROCC	0.7022	**0.8644**	0.7206	0.7246	**0.9066**	0.7946	**0.8824**	0.0209	0.8533	**0.9066**
	LCC	0.7745	**0.8805**	0.7912	0.7939	**0.9168**	0.8483	**0.897**	0.0794	0.8334	**0.9168**
	KROCC	0.5251	**0.6700**	0.5371	0.5423	**0.7442**	0.6058	**0.6959**	0.0666	0.67	**0.7442**
	RMSE	11.7999	**8.8423**	11.4075	11.343	**7.1366**	9.8785	**8.2474**	18.5956	10.1034	**7.1366**
	MAE	9.4311	**7.1463**	8.9226	8.8602	**5.7167**	7.7771	**6.7042**	15.2348	7.5061	**5.7167**

Table 3 (continued)

		FR IQA methods					NR IQA methods				
	Metric	SSIM	FSIM	SVQI	SQMS	PROPOSED	NIQE	ILNIQE	ASIQE	NRLT	PROPOSED
QACS	SROCC	0.8683	**0.9108**	**0.9194**	0.9096	0.8446	0.3413	0.2563	0.7005	**0.7048**	**0.8446**
	LCC	0.8696	**0.9073**	**0.9158**	0.9059	0.8552	0.4198	0.286	0.7046	**0.7435**	**0.8552**
	KROCC	0.691	0.7433	**0.7623**	**0.747**	0.6715	0.2583	0.1755	0.5182	**0.5471**	**0.6715**
	RMSE	1.0953	**0.933**	**0.8909**	0.9396	1.1319	2.0135	2.1259	1.5743	**1.4714**	**1.1319**
	MAE	0.8531	0.7183	**0.6608**	**0.6949**	0.8877	1.7139	1.8099	1.2281	**1.1669**	**0.8877**
SIQAD	SROCC	0.7566	0.5819	**0.8836**	**0.8803**	0.8417	0.3686	0.3197	0.757	**0.7911**	**0.8417**
	LCC	0.7615	0.5902	**0.8911**	**0.8872**	0.8633	0.3832	0.3866	0.7884	**0.8204**	**0.8633**
	KROCC	0.5682	0.425	**0.6985**	**0.6936**	0.6579	0.254	0.2471	0.5609	**0.6035**	**0.6579**
	RMSE	9.2784	11.5552	**6.4965**	**6.6039**	6.9270	13.2213	13.2012	8.8064	**8.0948**	**6.9270**
	MAE	7.1854	9.0116	**5.2282**	**5.2926**	5.5401	10.8467	10.7786	7.0635	**6.4199**	**5.5401**

Table 4 Average performance comparison on different databases

	SSIM	FSIM	SVQI	SQMS	NIQE	ILNIQE	ASIQE	NRLT	PROPOSED
Direct average performance									
SROCC	0.7583	**0.8228**	0.8035	0.8145	0.5677	0.5900	0.5062	0.7577	**0.8355**
LCC	0.7949	0.8299	0.8477	**0.8508**	0.6148	0.6317	0.5756	0.7807	**0.8602**
KROCC	0.5794	**0.6469**	0.6248	0.6328	0.4227	0.4442	0.3803	0.5849	**0.6679**
RMSE	7.0494	6.5225	**6.1005**	**6.1032**	8.6289	8.3724	10.1178	7.3756	6.1970
MAE	5.5377	5.2082	**4.8298**	**4.8195**	6.8756	6.7805	8.1333	5.6649	4.8417
Database size based average performance									
SROCC	0.7135	**0.7994**	0.7390	0.7746	0.4526	0.5356	0.5000	0.6968	**0.8040**
LCC	0.7536	**0.8285**	0.8097	0.8228	0.5107	0.5994	0.6242	0.7462	**0.8405**
KROCC	0.5387	**0.6284**	0.5670	0.5957	0.3274	0.3945	0.3766	0.5290	**0.6322**
RMSE	4.3005	4.1316	3.5870	**3.5843**	5.0741	4.9096	5.1880	3.8237	**3.3523**
MAE	3.3471	3.2257	2.8429	**2.8239**	4.0682	3.9774	4.1737	2.9628	**2.5985**

compared with ILNIQE and ASIQE. Third, the proposed method has advantages on the average performance across different databases, and the average SROCC and LCC values of the two calculation cases are both higher than 0.8. The average performance illustrated in Table 4 demonstrate the proposed method can be more suitable than existing NR IQA methods in the practical application scenarios that contain both natural image and SCI. However, the SCI is more complex than natural image which contain both document areas and graphic areas, thus, the accuracy of the NR IQA methods are lower than the FR IQA methods for fewer reference information.

In order to demonstrate the stability of the proposed method, the box plot of the distributions of SROCC values of different methods on each database is illustrated in Fig. 9. The ordinates are the values of SROCC, and the abscissas identify IQA methods. For each box, the central red line is the median value of the 100 train-test procedures on the certain database, the edges are the 25th and 75th percentiles, the whiskers extend to the most extreme data points not consider outliers. The results illustrated in Fig. 9 demonstrate the proposed method keeps stable on both natural image and SCI. Meanwhile, the proposed method gains more stable results on different databases. Therefore, the proposed method is suitable for the practical scenarios that image types and distortion types are varying.

Based on the above discussion, the proposed method can measure the quality accurately on both natural image and SCI. And, the proposed method keeps stable over different practical scenarios.

5 Conclusions

In this paper, a novel NR IQA method is proposed. Considering the limited performance of single IQA method. The proposed method is based on multiple quality prediction results. The proposed method first extract NSS feature, global structural feature and local structural feature. Relying on multiple features, the image can be described more comprehensive. Then, the quality prediction models are trained via different features, and the quality prediction results of different models are recorded. After that, the recorded quality scores are transformed to feature vectors, and SVR is employed to train the mapping relationship

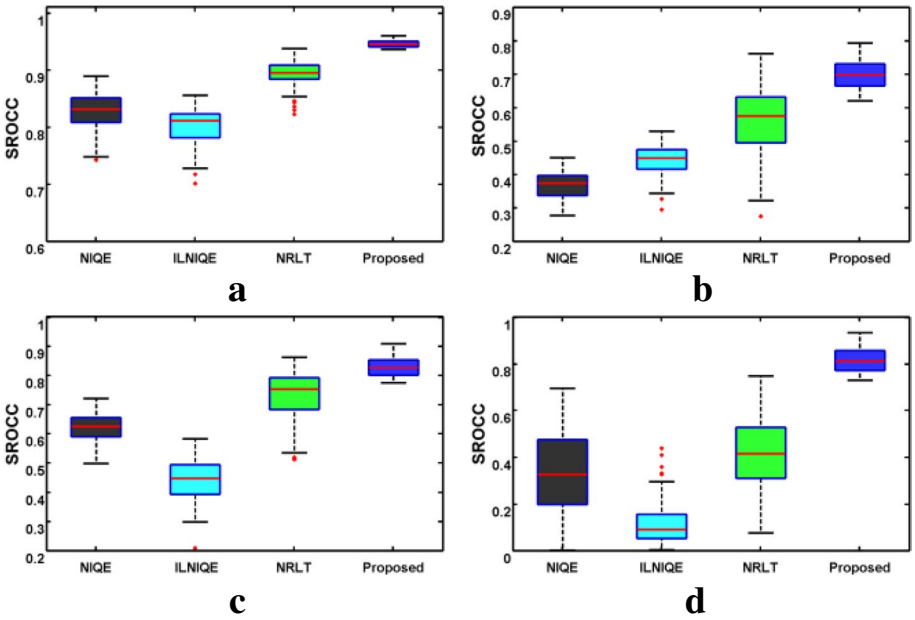

Fig. 9 Box plots of the distributions of SROCC values from 100 runs of different NR methods: (**a**) LIVE, (**b**) TID2013, (**c**) SIQAD, (**d**) QACS

between the feature vectors and subjective scores. Experimental results demonstrate the proposed method is effective for both natural image and SCI. Meanwhile, the proposed method keeps stable in different application scenarios, e.g. variation of the training set.

The performance of the histogram-based features that employed in the proposed method is limited, and more features can lead to more comprehensive description of the degradation in image. Therefore, our future work will focus on gain more quality-aware features to train more accurate and more generalized IQA model. In this way, the IQA method will better satisfy the complicated application scenarios in real life.

Acknowledgements This paper is supported by the Natural Science Foundation of the Jiangsu Higher Education Institutions of China (Grant No. 20KJB510021), the National Natural Science Foundation of China (Grant No. 62101268) and the National Natural Science Foundation of China (Grant No. 41971343).

Author contributions Conceived and designed the experiments: Xichen Yang. Performed the experiments: Xichen Yang. Analyzed the data: Xichen Yang. Wrote and reviewed the paper: Xichen yang, Genlin Ji, Tianshu Wang. Approved the final version of the paper: Xichen yang, Genlin Ji, Tianshu Wang.

Declarations

The authors have declared that there are no competing interests exist.

References

1. Arora M, Kumar M (2021) AutoFER: PCA and PSO based automatic facial emotion recognition. Multimed Tools Appl 80(2):3039–3049. https://doi.org/10.1007/s11042-020-09726-4
2. Bansal M, Kumar M, Kumar M (2020) 2D object recognition techniques: state-of-the-art work. Arch Comput Methods Eng. https://doi.org/10.1007/s11831-020-09409-1

3. Bansal M, Kumar M, Kumar M, Kumar K (2021) An efficient technique for object recognition using Shi-Tomasi corner detection algorithm. Soft Comput 25(6):4423–4432. doi: https://doi.org/10.1007/s00500-020-05453-y
4. Chang C-C, Lin C-J (2011) LIBSVM: A library for support vector machines. ACM Trans Intell Syst Technol 2(3):Article 27. https://doi.org/10.1145/1961189.1961199
5. Chhabra P, Garg NK, Kumar M (2020) Content-based image retrieval system using ORB and SIFT features. Neural Comput Appl 32(7):2725–2733. doi: https://doi.org/10.1007/s00521-018-3677-9
6. Dargan S, Kumar M, Ayyagari MR, Kumar G (2020) A survey of deep learning and its applications: a new paradigm to machine learning. Arch Comput Methods Eng 27(4):1071–1092. https://doi.org/10.1007/s11831-019-09344-w
7. Ding L, Huang H, Zang Y (2017) Image quality assessment using directional anisotropy structure measurement. IEEE Trans Image Process 26(4):1799–1809. https://doi.org/10.1109/TIP.2017.2665972
8. Fang Y, Yan J, Li L, Wu J, Lin W (2018) No reference quality assessment for screen content images with both local and global feature representation. IEEE Trans Image Process 27(4):1600–1610. https://doi.org/10.1109/TIP.2017.2781307
9. Garg D, Garg NK, Kumar M (2018) Underwater image enhancement using blending of CLAHE and percentile methodologies. Multimed Tools Appl 77(20):26545–26561. https://doi.org/10.1007/s11042-018-5878-8
10. Ghadiyaram D, Bovik AC (2016) Massive online crowdsourced study of subjective and objective picture quality. IEEE Trans Image Process 25(1):372–387. https://doi.org/10.1109/TIP.2015.2500021
11. Gu K, Wang S, Yang H, Lin W, Zhai G, Yang X et al (2016) Saliency-guided quality assessment of screen content images. IEEE Trans Multimed 18(6):1098–1110. https://doi.org/10.1109/TMM.2016.2547343
12. Gu K, Zhai G, Lin W, Yang X, Zhang W (2016) Learning a blind quality evaluation engine of screen content images. Neurocomputing 196:140–149. doi: https://doi.org/10.1016/j.neucom.2015.11.101
13. Gu K, Zhou J, Qiao J, Zhai G, Lin W, Bovik AC (2017) No-reference quality assessment of screen content pictures. IEEE Trans Image Process 26(8):4005–4018. https://doi.org/10.1109/TIP.2017.2711279
14. Gu K, Qiao J, Min X, Yue G, Lin W, Thalmann D (2018) Evaluating quality of screen content images via structural variation analysis. IEEE Trans Vis Comput Graph 24(10):2689–2701. https://doi.org/10.1109/TVCG.2017.2771284
15. Gu K, Qiao J, Lee S, Liu H, Lin W, Callet PL (2020) Multiscale natural scene statistical analysis for no-reference quality evaluation of DIBR-synthesized views. IEEE Trans Broadcast 66(1):127–139. https://doi.org/10.1109/TBC.2019.2906768
16. Gupta S, Mohan N, Kumar M (2020) A study on source device attribution using still images. Arch Comput Methods Eng. https://doi.org/10.1007/s11831-020-09452-y
17. Hu B, Li L, Wu J, Qian J (2020) Subjective and objective quality assessment for image restoration: A critical survey. Signal Process Image Commun 85:115839. https://doi.org/10.1016/j.image.2020.115839
18. Hui J, Chaoqiang L (2008) Motion blur identification from image gradients. 2008 IEEE Conference on Computer Vision and Pattern Recognition, p 1-8
19. Jabar F, Ascenso J, Queluz MP (2020) Objective assessment of perceived geometric distortions in viewport rendering of 360° images. IEEE J Selec Topics Signal Process 14(1):49–63. https://doi.org/10.1109/JSTSP.2019.2962970
20. Karaali A, Jung CR (2018) Edge-based defocus blur estimation with adaptive scale selection. IEEE Trans Image Process 27(3):1126–1137. https://doi.org/10.1109/TIP.2017.2771563
21. Ko H, Lee DY, Cho S, Bovik AC (2020) Quality prediction on deep generative images. IEEE Trans Image Process 29:5964–5979. https://doi.org/10.1109/TIP.2020.2987180
22. Kumar M, Chhabra P, Garg NK (2018) An efficient content based image retrieval system using BayesNet and K-NN. Multimed Tools Appl 77(16):21557–21570. https://doi.org/10.1007/s11042-017-5587-8
23. Kumar M, Jindal MK, Sharma RK, Jindal SR (2020) Performance evaluation of classifiers for the recognition of offline handwritten Gurmukhi characters and numerals: a study. Artif Intell Rev 53(3):2075–2097. doi: https://doi.org/10.1007/s10462-019-09727-2
24. Kumar M, Bansal M, Kumar M (2020) XGBoost: 2D-object recognition using shape descriptors and extreme gradient boosting classifier. International conference on Computational Methods and Data Engineering (ICMDE 2020)
25. Kumar A, Kumar M, Kaur A (2021) Face detection in still images under occlusion and non-uniform illumination. Multimed Tools Appl. https://doi.org/10.1007/s11042-020-10457-9

26. Lee D, Plataniotis KN (2016) Toward a no-reference image quality assessment using statistics of perceptual color descriptors. IEEE Trans Image Process 25(8):3875–3889. https://doi.org/10.1109/TIP.2016.2579308
27. Liao X, Li K, Yin J (2017) Separable data hiding in encrypted image based on compressive sensing and discrete fourier transform. Multimed Tools Appl 76(20):20739–20753. https://doi.org/10.1007/s11042-016-3971-4
28. Liao X, Yu Y, Li B, Li Z, Qin Z (2020) A New Payload Partition Strategy in Color Image Steganography. IEEE Trans Circuits Syst Video Technol 30(3):685–696. doi: https://doi.org/10.1109/TCSVT.2019.2896270
29. Liao X, Yin J, Chen M, Qin Z (2020) Adaptive payload distribution in multiple images steganography based on image texture features. IEEE Trans Dependable Secur Comput 1. https://doi.org/10.1109/TDSC.2020.3004708
30. Liu L, Wang T, Huang H (2019) Pre-attention and spatial dependency driven no-reference image quality assessment. IEEE Trans Multimed 21(9):2305–2318. https://doi.org/10.1109/TMM.2019.2900941
31. Liu Y, Gu K, Zhang Y, Li X, Zhai G, Zhao D et al (2020) Unsupervised blind image quality evaluation via statistical measurements of structure, naturalness, and perception. IEEE Trans Circuits Syst Video Technol 30(4):929–943. https://doi.org/10.1109/TCSVT.2019.2900472
32. Liu Y, Tang C, Zheng Z, Lin L (2020) No-reference stereoscopic image quality evaluator with segmented monocular features and perceptual binocular features. Neurocomputing 405:126–137. doi: https://doi.org/10.1016/j.neucom.2020.04.049
33. Mittal A, Moorthy AK, Bovik AC (2012) No-reference image quality assessment in the spatial domain. IEEE Trans Image Process 21(12):4695–4708. https://doi.org/10.1109/TIP.2012.2214050
34. Mittal A, Soundararajan R, Bovik AC (2013) Making a "Completely Blind" image quality analyzer. IEEE Signal Process Lett 20(3):209–212. https://doi.org/10.1109/LSP.2012.2227726
35. Moorthy AK, Bovik AC (2010) A two-step framework for constructing blind image quality indices. IEEE Signal Process Lett 17(5):513–516. https://doi.org/10.1109/LSP.2010.2043888
36. Ponomarenko N, Jin L, Ieremeiev O, Lukin V, Egiazarian K, Astola J et al (2015) Image database TID2013: Peculiarities, results and perspectives. Sig Process Image Commun 30:57–77. doi: https://doi.org/10.1016/j.image.2014.10.009
37. Qin M, Lv X, Chen X, Wang W (2017) Hybrid NSS features for no-reference image quality assessment. IET Image Process
38. Ruderman DL (1994) The statistics of natural images. Network 5(4):517-48. https://doi.org/10.1088/0954-898X_5_4_006
39. Sharifi K, Leon-Garcia A (1995) Estimation of shape parameter for generalized Gaussian distributions in subband decompositions of video. IEEE Trans Circuits Syst Video Technol 5(1):52–56. doi: https://doi.org/10.1109/76.350779
40. Sheikh HR, Sabir MF, Bovik AC (2006) A statistical evaluation of recent full reference image quality assessment algorithms. IEEE Trans Image Process 15(11):3440–3451. https://doi.org/10.1109/TIP.2006.881959
41. Sinno Z, Caramanis C, Bovik AC (2018) Towards a closed form second-order natural scene statistics model. IEEE Trans Image Process 27(7):3194–3209. https://doi.org/10.1109/TIP.2018.2817740
42. Tang L, Li L, Gu K, Sun X, Zhang J (2016) Blind quality index for camera images with natural scene statistics and patch-based sharpness assessment. J Vis Commun Image Represent 40:335–344. doi: https://doi.org/10.1016/j.jvcir.2016.07.007
43. VQEG (2000) Final report from the video quality experts group on the validation of objective models of video quality assessment
44. Wan Z, Gu K, Zhao D (2020) Reduced reference stereoscopic image quality assessment using sparse representation and natural scene statistics. IEEE Trans Multimed 22(8):2024–2037. https://doi.org/10.1109/TMM.2019.2950533
45. Wang S, Gu K, Zhang X, Lin W, Zhang L, Ma S et al (2016) Subjective and objective quality assessment of compressed screen content images. IEEE J Emerg Sel Top Circuits Syst 6(4):532–543. https://doi.org/10.1109/JETCAS.2016.2598756
46. Wu J, Zhang M, Li L, Dong W, Shi G, Lin W (2019) No-reference image quality assessment with visual pattern degradation. Inf Sci 504:487–500. doi: https://doi.org/10.1016/j.ins.2019.07.061
47. Yan B, Bare B, Tan W (2019) Naturalness-aware deep no-reference image quality assessment. IEEE Trans Multimed 21(10):2603–2615. https://doi.org/10.1109/TMM.2019.2904879
48. Yang H, Fang Y, Lin W (2015) Perceptual quality assessment of screen content images. IEEE Trans Image Process 24(11):4408–4421. https://doi.org/10.1109/TIP.2015.2465145
49. Yang X, Wang T, Ji G (2020) No-reference image quality assessment via structural information fluctuation. IET Image Proc 14(2):384–396. doi: https://doi.org/10.1049/iet-ipr.2019.0750

50. Yang X, Wang T, Ji G (2020) A local structural information representation method for image quality assessment. Multimed Tools Appl 79(31):22797–22823. https://doi.org/10.1007/s11042-020-09022-1
51. Yue G, Hou C, Gu K, Ling N, Li B (2018) Analysis of structural characteristics for quality assessment of multiply distorted images. IEEE Trans Multimed 20(10):2722–2732. https://doi.org/10.1109/TMM.2018.2807589
52. Yue G, Hou C, Zhou T (2019) Blind quality assessment of tone-mapped images considering colorfulness, naturalness, and structure. IEEE Trans Ind Electron 66(5):3784–3793. https://doi.org/10.1109/TIE.2018.2851984
53. Zhan Y, Zhang R, Wu Q (2017) A structural variation classification model for image quality assessment. IEEE Trans Multimed 19(8):1837–1847. https://doi.org/10.1109/TMM.2017.2689923
54. Zhang Y, Chandler DM (2018) Opinion-unaware blind quality assessment of multiply and singly distorted images via distortion parameter estimation. IEEE Trans Image Process 27(11):5433–5448. https://doi.org/10.1109/TIP.2018.2857413
55. Zhang L, Zhang L, Mou X, Zhang D (2011) FSIM: a feature similarity index for image quality assessment. IEEE Trans Image Process 20(8):2378–2386. https://doi.org/10.1109/TIP.2011.2109730
56. Zhang L, Zhang L, Bovik AC (2015) A feature-enriched completely blind image quality evaluator. IEEE Trans Image Process 24(8):2579–2591. https://doi.org/10.1109/TIP.2015.2426416
57. Zhou W, Bovik AC, Sheikh HR, Simoncelli EP (2004) Image quality assessment: from error visibility to structural similarity. IEEE Trans Image Process 13(4):600–612. doi: https://doi.org/10.1109/TIP.2003.819861
58. Zhou Z, Wang Y, Guo Y, Qi Y, Yu J (2020) Image quality improvement of hand-held ultrasound devices with a two-stage generative adversarial network. IEEE Trans Biomed Eng 67(1):298–311. https://doi.org/10.1109/TBME.2019.2912986