Frontiers of Information Technology & Electronic Engineering
www.jzus.zju.edu.cn; engineering.cae.cn; www.springerlink.com
ISSN 2095-9184 (print); ISSN 2095-9230 (online)
E-mail: jzus@zju.edu.cn

No-reference noisy image quality assessment incorporating features of entropy, gradient, and kurtosis*

Heng YAO[†1], Ben MA[2], Mian ZOU[2], Dong XU[†‡3,4], Jincao YAO[†‡3,4]

[1]*School of Optical-Electrical and Computer Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China*
[2]*School of Mechanical Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China*
[3]*Cancer Hospital of the University of Chinese Academy of Sciences (Zhejiang Cancer Hospital), Hangzhou 310000, China*
[4]*Institute of Basic Medicine and Cancer, Chinese Academy of Sciences, Hangzhou 310000, China*
[†]E-mail: hyao@usst.edu.cn; xudong@zjcc.org.cn; yaojc@zjcc.org.cn

Received Dec. 23, 2020; Revision accepted Mar. 4, 2021; Crosschecked June 8, 2021; Published online Oct. 11, 2021

Abstract: Noise is the most common type of image distortion affecting human visual perception. In this paper, we propose a no-reference image quality assessment (IQA) method for noisy images incorporating the features of entropy, gradient, and kurtosis. Specifically, image noise estimation is conducted in the discrete cosine transform domain based on skewness invariance. In the principal component analysis domain, kurtosis feature is obtained by statistically counting the significant differences between images with and without noise. In addition, both the consistency between the entropy and kurtosis features and the subjective scores are improved by combining them with the gradient coefficient. Support vector regression is applied to map all extracted features into an integrated scoring system. The proposed method is evaluated in three mainstream databases (i.e., LIVE, TID2013, and CSIQ), and the results demonstrate the superiority of the proposed method according to the Pearson linear correlation coefficient which is the most significant indicator in IQA.

Key words: Noisy image quality assessment; Noise estimation; Kurtosis; Human visual system; Support vector regression
https://doi.org/10.1631/FITEE.2000716 **CLC number:** TP753

1 Introduction

With the development of computer vision technology, image quality requirements have continuously been improved. Image quality assessment (IQA) has become the technical guidance for many image processing techniques, such as image denoising (Xu et al., 2020), compression (Li PY and Lo, 2018), hashing (Tang ZJ et al., 2018), and fusion (Guo et al., 2020). However, during the process of image acquisition, transmission, and display, many distortions occur in the image, including noise, blur, and blocking artifacts. Among them, noise contamination is the most common type of distortion.

In general, IQA can be divided into two aspects, subjective and objective. From reference information, existing objective quality metrics can be further categorized into full-reference (FR), reduced-reference (RR), and no-reference (NR) approaches. Because it is difficult to obtain reference image of the distorted image in general application scenarios, a series of NR-IQA metrics have been developed. Specifically, NR-IQA can be divided into two categories, general-purpose and distortion-specific methods. In general, a single type of distortion is more widespread in reality; for example, a noise-free image is often mixed with noise during the process of acquisition and transmission. Hence, the study of a high-precision IQA model

‡ Corresponding authors
* Project supported by the National Natural Science Foundation of China (No. 61702332) and the Zhejiang Provincial Natural Science Foundation of China (Nos. LZY21F030001 and LSD19H180001)
ORCID: Heng YAO, https://orcid.org/0000-0002-3784-4157; Dong XU, https://orcid.org/0000-0002-0583-240X; Jincao YAO, https://orcid.org/0000-0003-1543-6010

is beneficial to many denoising algorithms.

In practical applications, we have expanded on our previous work (Ma et al., 2020). In the previous work, the noise estimation algorithm designed by us was carried out in the transform domain. According to the central limit theorem and the independence assumption of noise, many types of non-Gaussian noise are similar to Gaussian noise in the transform domain. Thus, we can estimate many types of noise, including Gaussian noise, uniform noise, Laplacian noise, and so on. In this study, we design an objective NR-IQA method to evaluate the quality of noisy images based on noise estimation.

In the literature on general-purpose technique, an NR-IQA algorithm has been proposed based on the natural scene statistic (NSS) model that changes in the case of distortion to extract features in the wavelet coefficient transform domain, to identify distortion types and perceive image quality (Moorthy and Bovik, 2011). Mittal et al. (2012) and Saad et al. (2012) used the NSS model of the image discrete cosine transform (DCT) to extract features which were then used to predict image quality using the Bayesian inference model. Instead of operating in the transform domain, scholars proposed to extract 36 features derived from a spatial NSS model, and trained and evaluated the image quality using a support vector regression (SVR) based model (Mittal et al., 2012; Saad et al., 2012). An NR-IQA method for learning Gabor characteristics using unlabeled data available on the Internet was designed (Ye et al., 2012). Mittal et al. (2013) proposed to extract the statistical feature set of quality perception based on the spatial-domain NSS model to fit the multivariate Gaussian model. The quality was expressed as the distance between the extracted features of multivariate Gaussian fitting models from test and natural images. Then, Zhang et al. (2015) extracted NSS features to build a multivariate Gaussian model of the original natural images; then, this model was used as a reference to predict the quality of the image patches. The overall image quality score was finally obtained through average pooling (Zhang et al., 2015). Gu et al. (2015) extracted three groups of image features based on the free energy principle, the structure degradation model, and the human visual system (HVS). Then a support vector machine was used to train the regression model and score the image. Considering color influence on the image quality perception, a trained color dictionary was used to extract the structure and the color distortion feature map which were combined with brightness similarity to design an NR-IQA metric for color images (Li LD et al., 2016a). The characteristics of different color channels were compared to determine the optimal channel of the specific distortion of the image, followed by an image evaluation using color statistics (Wang Q et al., 2016). A blind image blur assessment metric based on discrete Tchebichef moments was proposed by normalizing the moment energy using the variance of gradient image blocks (Li LD et al., 2016b). Quaternion wavelet transform was used to depict the intrinsic structures of the image and decompose the image with three scales, and a prediction model was designed using a random forest (Tang LJ et al., 2017). This method used the Gaussian derivative filter to decompose the image by gradient magnitude and phase to obtain local statistical features. These features were then used to predict the quality score via SVR (Zhou et al., 2017). Li QH et al. (2017) used two complementary information features from the feature image and texture image using Prewitt linear filters and the local contrast normalization. Then, SVR was applied to map the feature space into the human perception score (Li QH et al., 2017). Nonlinear features of images were extracted using higher-dimensional mathematical expressions and were applied to the IQA framework. At present, they were applied only to the FR evaluation algorithm (Ding et al., 2016). An assessment method was introduced using a deep convolutional neural network (CNN) to learn features and natural scene statistics through data driven methods (Bosse et al., 2016). An NR-IQA method based on a modified CNN structure was proposed (Pan et al., 2016). Each image patch was scored by a CNN, and a saliency map was computed using free energy based on the neural theory. Finally, the evaluation of the whole image was completed considering the saliency map to set different weights for different patches (Pan et al., 2016). Multi-scale and multi-resolution quality features, including sharpness-aware features and entropy features, were extracted in both the spatial and spectral domains to learn a sharpness quality model (Li LD et al., 2017). Min et al. (2018) used the FR-IQA framework to design an NR-IQA method by measuring the similarities between the distortion image and multiple

pseudo-reference images. An orientation similarity based pattern (OSP) was extracted via the rotation invariant and pattern similarity, and the feasibility of OSPs was discussed as a feature after establishing the visual pattern degradation model to evaluate the images (Wu et al., 2019). Based on the principle of free energy and the human eye motion mechanism, a metric using a trained attention-driven model extracted quality maps, distortion type maps, and related weight maps. Region-of-interest pooling and weighted averaging were employed to integrate all estimated maps of the salient region into the overall predicted values (Chen et al., 2020). Based on deep meta-learning, bilevel gradient optimization was used to learn the shared prior knowledge model of various distortions to obtain the target quality model (Zhu HC et al., 2020). A feature-segmentation strategy was adopted to design a CNN model including five convolutional layers with max-pooling, one special fully-connected layer with feature segmentation, and one output layer (Shen et al., 2020). Recently, a novel NR-IQA model was proposed for screen content images considering the large difference that exists in different regions and using noise classification tasks and Siamese networks (Jiang et al., 2020). Given the lag in quality evaluation of image restoration and the weak robustness of evaluation indicators, a survey was presented to deeply analyze the differences between traditional image quality assessment methods and image restoration quality assessment methods (Hu et al., 2020).

To further study the changes in human visual quality caused by noise distortion and conduct a reasonable evaluation, a series of NR-IQA algorithms dedicated to noisy images have been proposed. Kong et al. (2013) maximized the consideration of two groups of structural similarity maps, one from the noisy input image and the estimated noise signal and the other from the noisy input image and the denoised image. The linear correlation coefficient of the two response maps was taken as the predicted value (Kong et al., 2013). Then, a method was proposed to estimate noise variance based on image block homogeneity combined with the masking effect of the HVS (Huang et al., 2014). Liu et al. (2014) predicted the image quality based on the free energy brain principle, image gradient extraction, and texture masking. This is similar to the human visual perception of noise. An NR-IQA method was designed using the perception-weighted probability summation model of local noise, to predict the perceived relative amount of noise in images with different contents (Zhu T and Karam, 2014). Zhai et al. (2015) designed a dual-model (i.e., near-threshold and super-threshold models) prediction method based on the noise estimation of images (Zhai and Wu, 2011) and the neural and psychological characteristics of the HVS (Zhai et al., 2012). Gradient singular value decomposition (GSVD) was used to obtain the features, and the Gaussian mixture model was employed to obtain the feature space model (Ospina-Borras and Restrepo, 2016). In the feature space, the Euclidean distance between the reference model and the assessed image was used as the prediction score (Ospina-Borras and Restrepo, 2016). Buczkowski (2018) proposed an average opinion score prediction method based on noise estimation. It is suitable for a nonsingle image database. Then, the sensitivity of the HVS to local image distortion was considered, and a visual significance model was established of image patches based on the method of using the adaptive visual significance index to evaluate the image (Oszust, 2019). Recently, Deng et al. (2020) studied the validity of the kurtosis of coefficient samples obtained by the image after wavelet transform as a noise distortion feature, and an extreme learning machine was applied to train the prediction model. This can map kurtosis values into perceptual quality scores and can evaluate five types of noisy image (Deng et al., 2020).

Through the study, we found that most prediction methods (Mittal et al., 2012; Deng et al., 2020) were designed in the same domain (e.g., the spatial, wavelet, or gradient domain), whereas few methods were involved operations in multiple domains and were explored the mutual relationship between different domains. Therefore, we comprehensively extract features from three different domains to establish a model that can map all features into a quality score. The main contributions of this paper can be summarized as follows:

1. The image entropy feature is extracted based on the combination of noise estimation in the frequency domain and variation coefficient in the gradient domain.

2. The rationality of kurtosis in the principal component analysis (PCA) domain as human-

perceived noise distortion is thoroughly analyzed, and an adaptive optimization of the kurtosis feature in combination with the visual masking effect is conducted.

3. In the process of SVR training of the regression model, the particle swarm optimization (PSO) algorithm is used to improve the prediction accuracy of the model.

2 Image high-order statistics and the application in noise variance estimation

At present, most noisy image assessment algorithms are based on NSS characteristics. Because noise signal is the main factor that interferes with the visual quality of human perception, it is feasible to estimate the noise level for the evaluation of noisy images, such as the algorithms designed by Huang et al. (2014) and Zhai et al. (2015), and the evaluation index is significantly improved compared with the previous methods. However, we find that even for images with the same noise level, the subjective quality scores, i.e., mean opinion score (MOS) and differential MOS (DMOS), still have significant differences.

An example of these differences is illustrated in Fig. 1, where all 30 images from the categorical subjective image quality (CSIQ) database (Larson and Chandler, 2010) with the same noise level of 2 are involved to compare the subjective scores. It is not difficult to observe that the subjective scores are unstable even though they have the same noise distortion level. Based on this phenomenon, even with an accurate estimate of noise levels, estimated noise levels cannot be directly used to evaluate the image quality. Therefore, we design an NR-IQA algorithm to evaluate the noisy image quality based on the feature combination of entropy, gradient, and kurtosis.

Because noise estimation is the basis of noisy image assessment, we employ the noise estimation method based on the invariance of high-order statistics, which has been proposed in Ma et al. (2020). Before reviewing this method, we first briefly discuss the correlation expression of higher-order statistics (e.g., kurtosis and skewness) and the related characteristics. For a random variable z, skewness $s(z)$ and kurtosis $\kappa(z)$ are defined as follows:

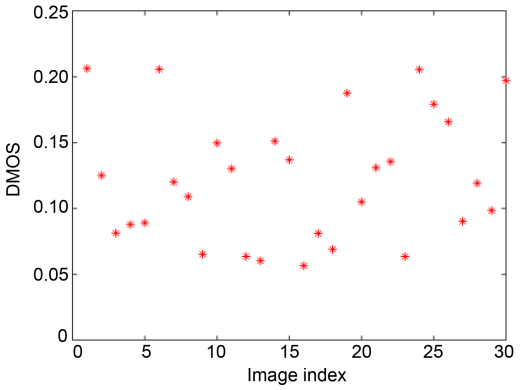


Fig. 1 Subjective scores of all 30 images from the CSIQ database with the same noise distortion level of 2

$$\begin{cases} s(z)=\dfrac{\mu_3(z)}{\sigma^3(z)}, \\ \kappa(z)=\dfrac{\mu_4(z)}{\sigma^4(z)}, \end{cases} \tag{1}$$

where $\mu_\eta(z)$ and $\sigma^\eta(z)$ (here, η=3 or 4) represent the η^{th} central moment and the η^{th} power of the standard deviation of z, respectively. Specifically, $\mu_\eta(z)$ and $\sigma^\eta(z)$ can be computed as follows:

$$\begin{cases} \mu_\eta(z)=E\left([z-E(z)]^\eta\right), \\ \sigma^\eta(z)=\{\sigma^2(z)\}^{\eta/2}=\left\{E\left([z-E(z)]^2\right)\right\}^{\eta/2}, \end{cases} \tag{2}$$

where $E(\cdot)$ represents the expectation operator.

Zoran and Weiss (2009) first proposed to transform a clean image into the frequency domain through specific filters. The kurtosis has the invariable characteristic, and the experimental results indicated that the factor causing the fluctuation in this invariability is the noise in the clean image. Based on this observation, many high-performance methods have been designed for noise estimation (Lyu et al., 2014; Dong et al., 2017) and IQA metrics (Deng et al., 2020). Recently, we proposed a noise estimation algorithm based on the skewness invariance and adapted the noise injection (Ma et al., 2020). Assuming that X is a clean image and that N is an independent additive white Gaussian noise (AWGN) signal, image Y polluted by noise conforms to the following:

$$Y=X+N. \tag{3}$$

The noisy image is transformed by DCT filters, and its expression in the frequency domain can be modeled as follows:

$$y=x+n, \tag{4}$$

where y, x, and n represent the noisy image, clean image, and independent noise in the transform domain, respectively. According to the relationship between the statistics and the independence of noise, we transform the problem of estimating the noise level to a problem of a nonlinear optimization function with specific constraints:

$$\begin{aligned}&\left[\hat{s}(x),\hat{\sigma}^2(n)\right]=\\&\arg\min_{s_x,\sigma_n^2}\sum_{k=2}^{8^2}\left[s(y_k)-\left(\frac{\sigma^2(y_k)-\sigma^2(n)}{\sigma^2(y_k)}\right)^{\frac{3}{2}}s(x)\right]^2\\&\text{subject to }\frac{1}{\left|8^2-1\right|}\sum_{k=2}^{8^2}\left|\hat{s}(y_k)\right|\le\left|s(x)\right|,\end{aligned} \tag{5}$$

where $\sigma^2(y_k)$, $s(y_k)$, and $\hat{s}(y_k)$ represent the variance and the actual and the estimated skewness of a noisy image in the k^{th} component, respectively, $s(x)$ and $\hat{s}(x)$ represent the actual and the estimated skewness of a clean image, respectively, and $\sigma^2(n)$ and $\hat{\sigma}^2(n)$ represent the actual and the estimated noise variances, respectively. In the DCT, an 8×8 DCT filter is selected. Noise usually exists in the alternating current frequency components; therefore, the components from 2 to 8^2 are involved in the estimation, as shown in Eq. (5). Additionally, an adaptive noise injection strategy is applied in the noise estimation algorithm. According to the image complexity, the threshold model of the noise level can be distinguished. By comparing the noise threshold value with the noise estimation result, we can judge whether the noise injection should be carried out for the second round estimation. Finally, the noise level of the image can be determined and is still denoted as $\hat{\sigma}(n)$. For details, readers can refer to Ma et al. (2020).

3 Proposed method for NR-IQA

Because noise affects the human visual quality in a noisy image, in this section, we describe our method for extracting features in different domains. Fig. 2 illustrates the flow diagram of the proposed method. For a given noisy image, three-domain decompositions are first constructed, including DCT, gradient, and PCA domains. We first estimate the noise level $\hat{\sigma}(n)$ of the given test image based on the scale invariance of skewness in the DCT domain and adapt the noise injection strategy, as introduced in Section 2. Then, the estimated noise level is used to calculate the differential entropy of the noisy image. Considering the HVS, the coefficient of variation that reflects the masking effect is applied to adaptively optimize the proposed entropy as a distortion measure feature of the test image. The second proposed feature is built into the gradient domain. The second proposed feature uses four directional filters to generate gradient maps that are comprehensively considered to obtain the gradient feature. The third feature of the kurtosis feature is designed using the assumption of scale invariance of kurtosis in the PCA domain and combining the effect of the image content. Finally, the quality model using SVR training is employed to predict the objective score of the test image. Detailed descriptions of the proposed method are presented in the following subsections.

3.1 Extraction of gradient and entropy features

For a noisy image, it is meaningful to estimate the standard deviation to evaluate the image; however, in some cases, the noise standard deviation cannot completely measure the quality. Table 1 lists the comparison between different images from the CSIQ database with respect to the noise level, the subjective quality score (i.e., DMOS value), and the image complexity, where the image complexity is computed using the gradient information of all pixels. Specifically, as in our previous work (Ma et al., 2020), we select the maximum gradient of each pixel in the four different directions and then average the maximum gradient values of all pixels to obtain the image complexity of the images.

Table 1 indicates that the DMOSs of different images are different although they have the same noise level, and that these differences are related to the image complexity. It is even more significant for images 21 and 26. The masking effect of human vision is affected by the texture information and noise interference in the image. In our method, the entropy

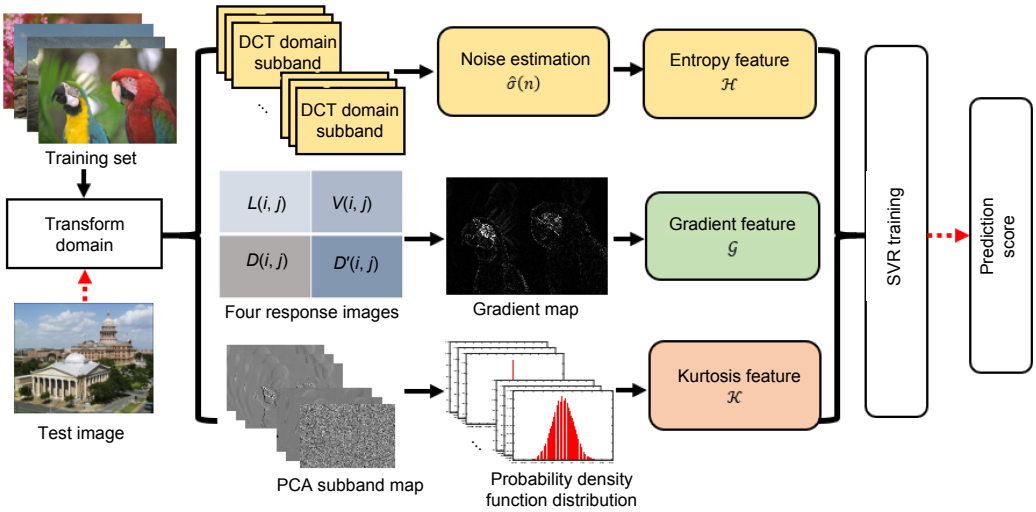


Fig. 2 Flow diagram of the proposed method

characteristics of each noisy image are extracted by combining the noise level and the visual masking effect. Suppose that the noise obeys a Gaussian distribution with a mean of zero and a standard deviation of σ, (i.e., $N\sim\mathcal{N}(0, \sigma)$). Then its probability density function is expressed as follows:

$$f(N)=\frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{N^2}{2\sigma^2}\right). \tag{6}$$

The differential entropy of image noise denoted as φ_N can be expressed as follows:

$$\varphi_N=-\int f(N)\ln f(N)\,\mathrm{d}N. \tag{7}$$

Based on Eqs. (6) and (7), converting the unit of the logarithm function to bit, φ_N can be further expressed as follows:

$$\varphi_N=\log_2\sqrt{2\pi\mathrm{e}\sigma^2}. \tag{8}$$

Therefore, the entropy calculation of the given noisy image with the estimated noise standard deviation $\hat{\sigma}(N)$ which is interpreted in Section 2 can be executed as

$$\varphi_N=\log_2\sqrt{2\pi\mathrm{e}\hat{\sigma}^2(N)}. \tag{9}$$

To compensate for the difference in human vision caused by the inherent information content of the image in the process of information entropy calculation, we introduce a variation coefficient of the image complexity based on the gradient calculation to obtain the final information entropy feature, denoted as $\mathcal{H}$. For any given image pixel $Y(i,j)$, where (i,j) denotes the coordinates of the pixel, its gradients along different directions for different filters can be calculated. Fig. 3 depicts the convolutional filters applied in our method, where the masks are denoted as F_L, F_V, F_D, and $F_{D'}$. Next, four different direction filters are convolved with image Y to obtain four different gradient maps, and the corresponding convolution values of $Y(i, j)$ are denoted as $L(i, j)$, $V(i, j)$, $D(i, j)$, and $D'(i,j)$:

$$\begin{cases}L(i,j)=Y(i,j)\otimes F_L,\\ V(i,j)=Y(i,j)\otimes F_V,\\ D(i,j)=Y(i,j)\otimes F_D,\\ D'(i,j)=Y(i,j)\otimes F_{D'},\end{cases} \tag{10}$$

where "⊗" represents the convolution operation.

Table 1 Comparison of different images with respect to the noise level, DMOS, and image complexity in the CSIQ database

Image index	Noise standard deviation level	DMOS	Image complexity
21	1	0.0266	12.76
91	1	0.0509	21.77
26	1	0.1043	26.75

−1	−1	−1
0	0	0
1	1	1

F_L

−1	0	1
−1	0	1
−1	0	1

F_V

0	1	1
−1	0	1
−1	−1	0

F_D

−1	−1	0
−1	0	1
0	1	1

$F_{D'}$

Fig. 3 Gradient filters in four different directions

Then, the gradient values along four directions of the same pixel are compared, and the smallest value is determined as the minimum gradient of the pixel, denoted as $G(i, j)$. The average gradient value of the entire noisy image, denoted as $\overline{g}$, is calculated as the gradient feature:

$$\overline{g}=\frac{\sum_{i=1}^{R}\sum_{j=1}^{C}G(i,j)}{RC}, \tag{11}$$

where R and C represent the numbers of row and column of the minimum gradient map, respectively.

According to the image gradient map, the dispersion degree of the image gradient δ is defined as the ratio of the standard deviation of the gradient image $G(i, j)$ to the average gradient value $\overline{g}$, and is expressed as

$$\delta=\frac{1}{\overline{g}}\sqrt{\frac{\sum_{i=1}^{R}\sum_{j=1}^{C}\left(G(i,j)-\overline{g}\right)^2}{RC}}. \tag{12}$$

The reason for introducing coefficient δ is to offset the difference in visual quality caused by different image complexities with the same noise level for various test images. Through δ, the entropy feature of the image can be expressed as follows:

$$\mathcal{H}=\varphi_N/\delta. \tag{13}$$

In addition, as seen through experiments, the change in the image gradient is closely related to the influence of the noise; thus, the average gradient $\overline{g}$ is also an effectual feature, and this feature is denoted as $\mathcal{G}$. Fig. 4 illustrates several examples of the minimum gradients of some standard test images; Figs. 4a–4c show the selected images, Figs. 4d–4f indicate their corresponding visual gradient maps, and Figs. 4g–4i present the corresponding mean gradient values with different noise levels. As displayed in Figs. 4g–4i, with the increase in the image distortion degree, the gradient value of the image also increases; this phenomenon demonstrates that gradient feature $\mathcal{G}$ is monotonic to the noise level. Similarly, as shown in Figs. 4j–4l, we can see that the entropy feature also has obvious monotony, which is taken as an important feature in improving the performance of the algorithm in this study.

3.2 Extraction of kurtosis feature

Recently, a noisy IQA method was designed based on the hypothesis proposed by Zoran and Weiss (2009) and the advantages and disadvantages of kurtosis characteristics in DCT, random unitary transform, and discrete wavelet transform were compared (Deng et al., 2020). In Lyu et al. (2014), kurtosis differences of natural images were compared in various linear transform domains in depth, and the experimental and theoretical proofs for the hypothesis were provided. To make the extracted features fully reflect the degree of distortion in the image, we choose PCA as a tool for linear transformation and further improve the kurtosis characteristics. To obtain the kurtosis value in the frequency domain, the direct current (DC) component of the input image Y is first removed as follows:

$$Y'=Y-E(Y), \tag{14}$$

where Y' represents the noisy image without the DC component.

Then, Y' is convolved with a two-dimensional band-pass filter P_m selected from a $d\times d$ linear transform base, where m is the index for all band-pass filters and $1\leq m\leq d^2-1$. The response image R_m can be obtained as follows:

$$R_m=Y'\otimes P_m. \tag{15}$$

The response image R_m is first transformed into a column vector, and the kurtosis of each response image is obtained using Eqs. (1) and (2), denoted as κ_m. All obtained response kurtosis values κ_m are then averaged as follows:

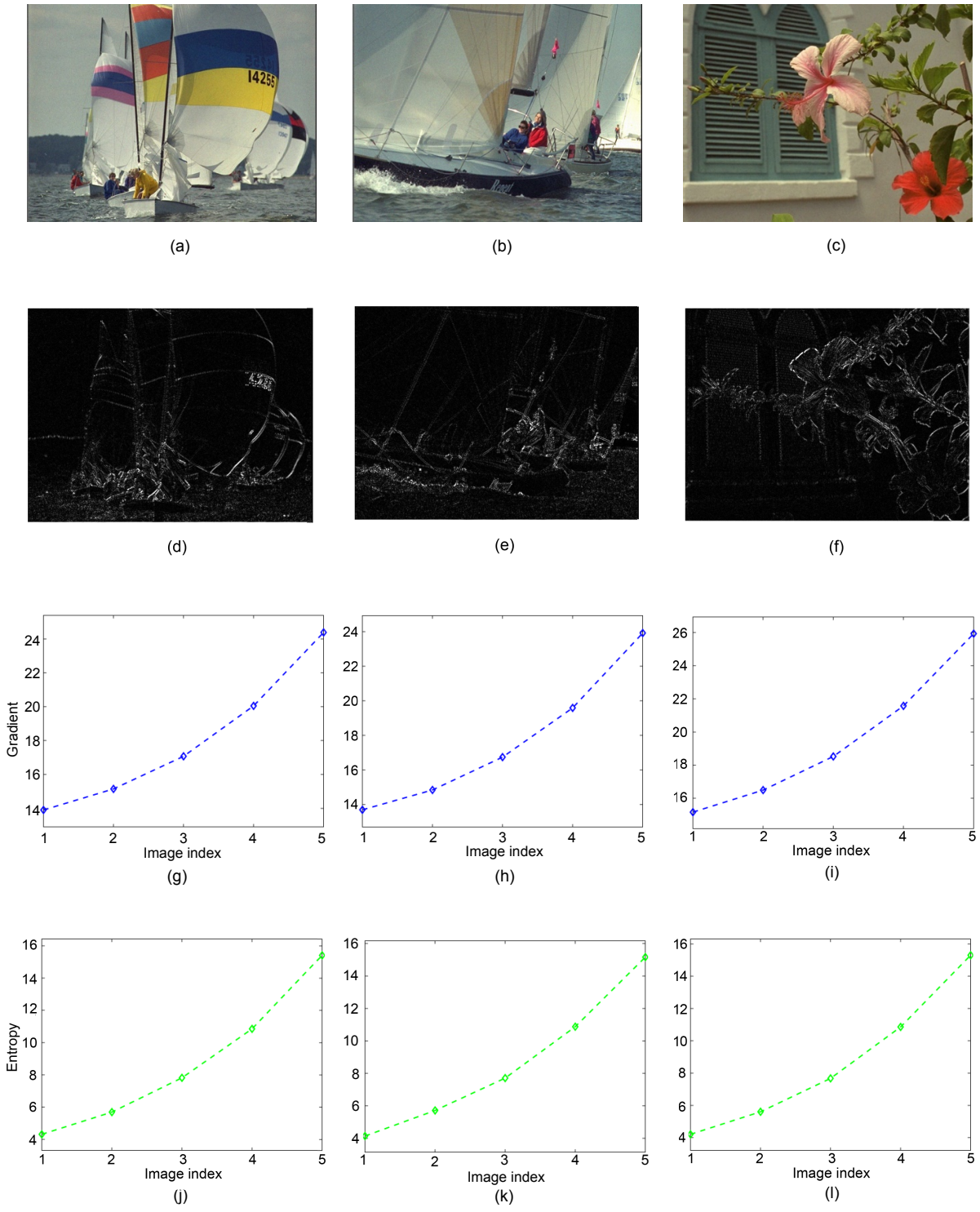


Fig. 4 Examples of gradient maps and gradient trends with different noise levels

The first row (a–c) shows the noisy images from the Tampere image database 2013 (TID2013) (Ponomarenko et al., 2013). The second (d–f), third (g–i), and fourth (j–l) rows are the gradient maps, mean gradient values, and entropy feature values of the noisy images, respectively. In each subfigure, in the third and fourth rows (g–l), five diamonds from left to right correspond to the same image with different noise levels ranging from low to high

$$\bar{\kappa}=\frac{\sum \kappa_m}{d^2-1}, \tag{16}$$

where $\bar{\kappa}$ denotes the mean kurtosis of the noisy images. To verify whether the kurtosis statistics satisfies the requirements of a noisy image feature, we list several examples of kurtosis under different images with Gaussian noise distortion, as illustrated in Fig. 5. We randomly select four noisy images, as shown in Figs. 5a–5d, and according to the PCA filters, we transform the images to the frequency domain and draw the coefficient distribution of the response images. As presented in Figs. 5e–5h, the coefficient distributions in the transform domain are different from the distortion degree. Specifically, compared with the blue dashed curves (images with a lower subjective quality score), the red solid curves (images with a higher subjective quality score) are more leptokurtic. To fully reveal the effect of the distortion on kurtosis, we measure the Gaussian noise distortion with different severity degrees for each noisy image, as depicted in Figs. 5i–5l. With the reduction in the DMOS, the average kurtosis $\bar{\kappa}$ gradually decreases and exhibits obvious monotonicity, verifying the ability of the average kurtosis calculated using PCA basis filters to represent the quality assessment score to a sufficient extent.

As mentioned in Dong et al. (2017), we observe an important phenomenon (Fig. 6). In the entire image with both complex and smooth areas, as indicated in Fig. 6a, the kurtosis consistency at different frequencies is not stable. When the image is divided into complex and smooth areas, as presented in Fig. 6c, the kurtosis values of the two parts are not consistent.

Specifically, kurtosis consistency in the smooth area performs well, whereas that in the complex area performs unstably. This phenomenon requires us to consider the influence of the image content on the calculation of $\bar{\kappa}$. Therefore, in our method we use δ which was designed in Section 3.1 to reduce the influence of the image content on the kurtosis calculation. Therefore, the final kurtosis feature $\mathcal{K}$ is defined as follows:

$$\mathcal{K}=\frac{\bar{\kappa}}{\delta}. \tag{17}$$

3.3 Model training and score prediction

In this work, three features are extracted for each image, information entropy feature $\mathcal{H}$, gradient feature $\mathcal{G}$, and kurtosis feature $\mathcal{K}$. To combine them as a whole to evaluate the image quality, we apply SVR (Chang and Lin, 2011) to train the quality prediction model. Then, the trained SVR model is used to score the noisy images. In the implementation, the radial basis function (RBF) is used as the SVR kernel. We apply the PSO (Kennedy and Eberhart, 1995) to determine the optimal model parameters of penalty coefficient c and parameter coefficient of RBF γ during the training process. The steps are as follows:

1. Set the ranges of variables c and γ of the training model. The position and speed ranges of c are set to [0.1, 100] and [−60, 60], respectively, and the initial position and size range of γ are set at [0.1, 1000] and [−600, 600], respectively.

2. The particle swarm size is set at 20, and the iteration number is set at 100.

The particle position and velocity are randomly initialized, and the fitness of the particle is calculated according to the fitness function. Through loop iteration, the position and velocity of the particles are updated continuously, and the extreme values of individuals and groups are updated according to the fitness value of the new particles. After the iteration, we draw the curve of the fitness value for each generation, as illustrated in Fig. 7. The fitness tends to be stable as the iteration number increases.

4 Experimental results and analysis

To quantitatively evaluate the performance of the proposed method, we choose two common criteria recommended by the Video Quality Experts Group (2003), i.e., the Pearson linear correlation coefficient (PLCC) and the Spearman rank-order correlation coefficient (SRCC) (Sheikh et al., 2006). Specifically, PLCC and SRCC are usually used to measure the accuracy and monotonicity of the results, respectively, and the values closer to 1 indicate a better performance. The relationship between the subjective and the predicted quality scores may be nonlinear because of nonlinear responses of human observers.

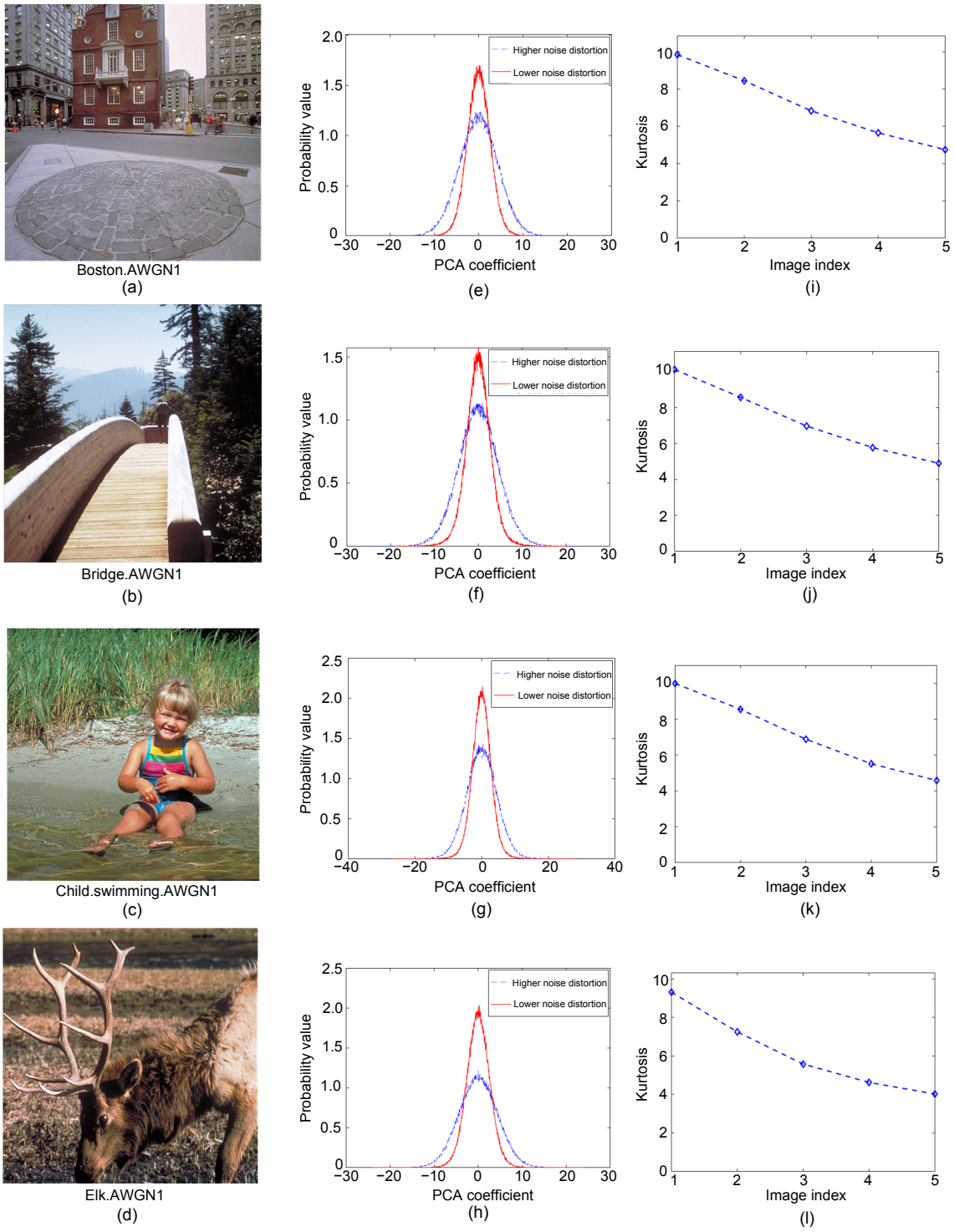


Fig. 5 Examples of probability density distributions and kurtosis trends with noise levels

The left column (a–d) presents the randomly selected images from the CSIQ image database. The middle column (e–h) displays a comparison of the coefficient distribution of the same image with different distortion degrees at a specific component R_{63} in the transform domain, where the blue dashed and the red solid curves represent the response image coefficient distributions with higher and lower noise distortion degrees, respectively. The right column (i–l) reveals the changes of the corresponding kurtosis $\bar{\kappa}$ with an increase in the subjective score (i.e., DMOS value). Five diamonds in subfigures (i)–(l) from left to right are with different noise levels ranging from low to high. References to color refer to the online version of this figure

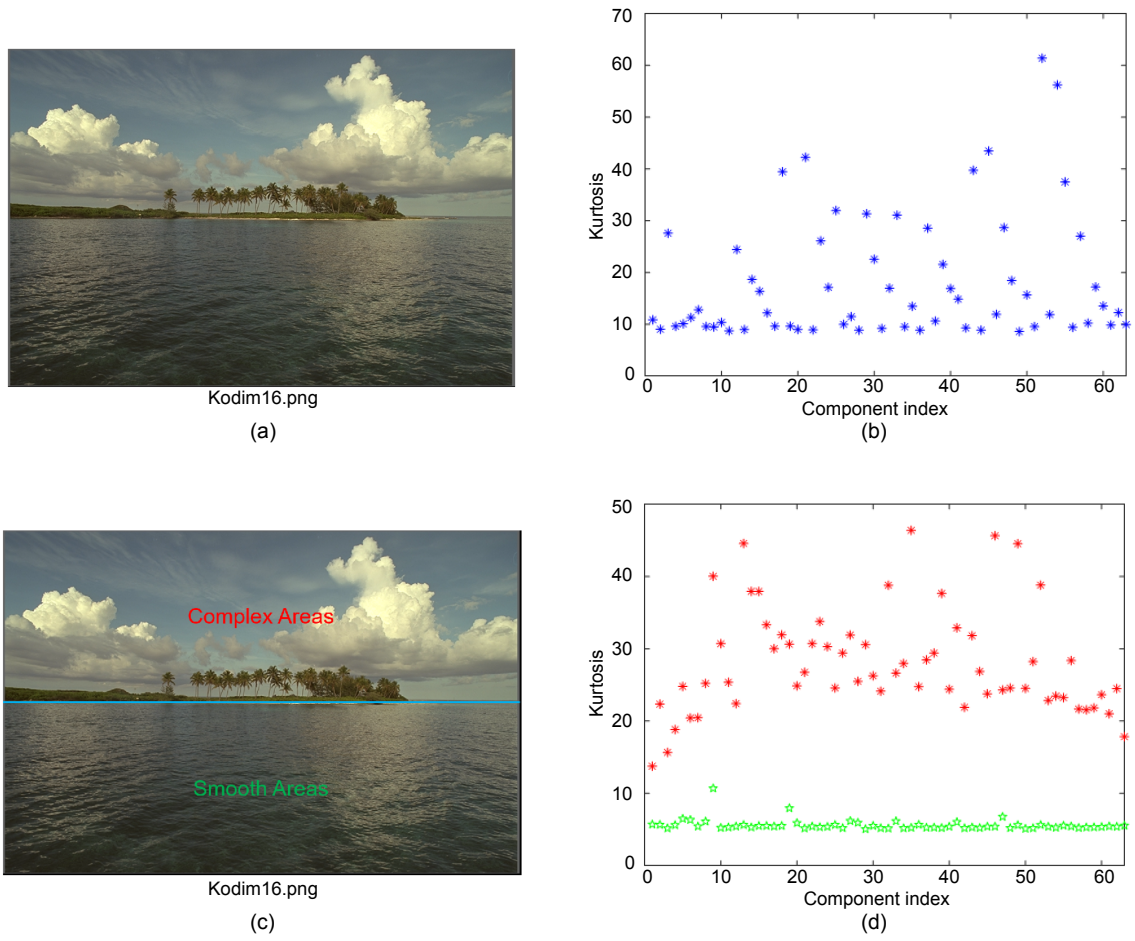


Fig. 6 Comparison of the scale invariance of the entire image and the segmented image

(a) and (c) are the images from the laboratory for image & video engineering (LIVE) database (http://live.ece.utexas.edu/research/quality/) and its segmented image, respectively. (b) shows the kurtosis values of (a) at different component indices. (d) shows the kurtosis values of the segmented image areas at different component indices, where the red asterisks and the green pentagons represent the complex and smooth areas of (c), respectively. References to color refer to the online version of this figure

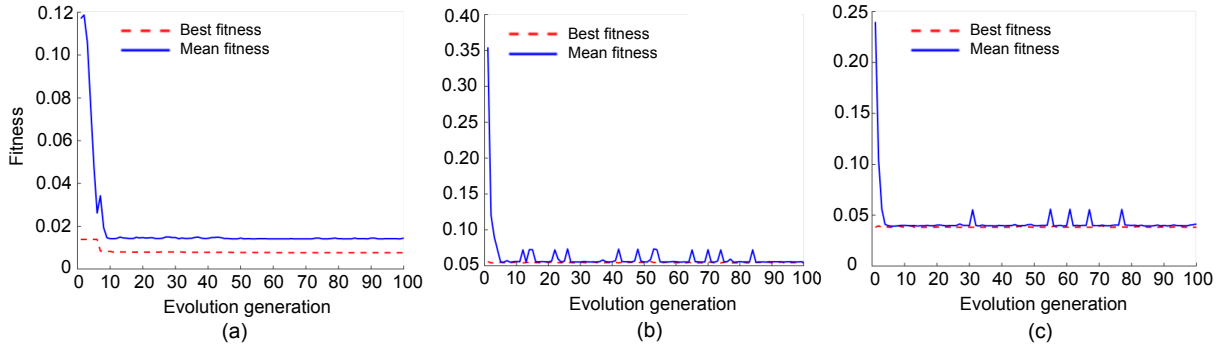


Fig. 7 Fitness curves for the particle swarm optimization (PSO): (a) LIVE; (b) TID2013; (c) CSIQ

Before calculating these values, the following five-parameter logistic regression function is defined (Sheikh et al., 2006):

$$\xi(p)=\alpha_1\left(\frac{1}{2}-\frac{1}{1+\mathrm{e}^{\alpha_2(p-\alpha_3)}}\right)+\alpha_4 p+\alpha_5, \quad (18)$$

where α_1–α_5 are the fitting model parameters, and p and $\xi(p)$ are the predicted scores before and after the regression process, respectively. Here, we comprehensively measure the performance of the proposed method in the public image databases, LIVE, TID2013, and CSIQ. Detailed information on the three benchmark databases is listed in Table 2, and some test images are shown in Fig. 8.

We compare our proposed model against representative state-of-the-art NR-IQA models, including DIIVINE (Moorthy and Bovik, 2011), BLIINDS (Saad et al., 2012), BRISQUE (Mittal et al., 2012), CORNIA (Ye et al., 2012), NIQE (Mittal et al., 2013), PWN (Zhu T and Karam, 2014), DMDM (Zhai et al., 2015), GSVD (Ospina-Borras and Restrepo, 2016), QENI (Oszust, 2019), and BNIQAK (Deng et al., 2020), where the methods DIIVINE, BLIINDS, BRISQUE, CORNIA, and NIQE are the general-purpose IQA methods, and PWN, DMDM, GSVD, QENI, and BNIQAK are the noise-specific methods. In addition, two FR-IQA methods, i.e., peak signal-to-noise ratio (PSNR) and structural similarity index measure (SSIM) (Wang Z et al., 2004), are involved

Table 2 Three benchmark databases for the IQA

Database	Number of reference images	Distortion type	Distortion level	Number of distortion images	Subjective score
LIVE	29	AWGN	5	145	DMOS
TID2013	25	AWGN	5	125	MOS
CSIQ	30	AWGN	5	150	DMOS

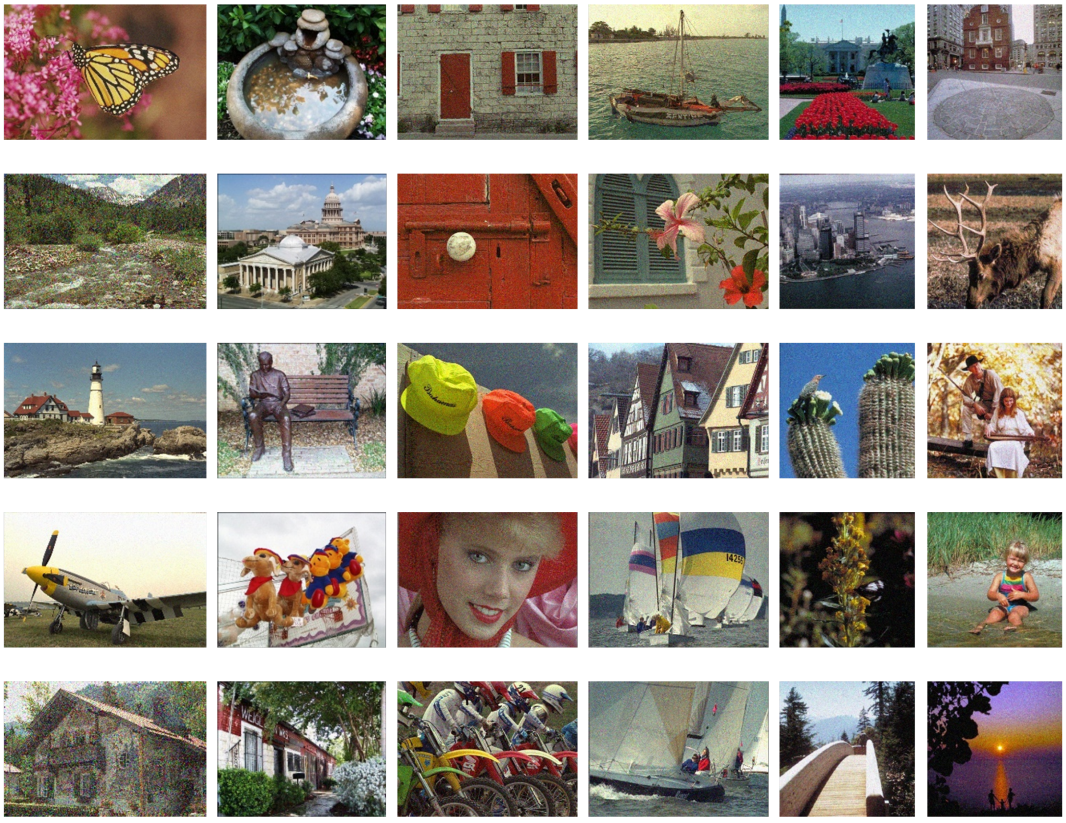


Fig. 8 Partial test noisy images from the three different databases

The first and second columns are obtained from the LIVE database, the third and fourth columns are obtained from the TID2013 database, and the fifth and sixth columns are obtained from the CSIQ database

in the comparison. Table 3 lists the experimental results, and the rank of each method is also listed in parentheses after each result. Eighty percent images are randomly selected from each image database to train the prediction model, and the remaining 20% images are used to verify the prediction results. To avoid bias, the training test is repeated 1000 times, and the median value of the 1000 results is taken as the performance of the algorithm. Table 3 indicates that in addition to the criterion of the SRCC of the PSNR in the LIVE image database, our proposed method and the BNIQAK reach the top two in the all comparison methods and exceed all comparison NR-IQA methods. The average PLCC rank of our method is 1.33, which is the best result among those of the methods.

Specifically, for the FR-IQA methods, the PSNR exhibits an excellent evaluation performance, reaching more than 0.9 on two criteria, whereas the SSIM does not perform well on the TID2013 database. We can see that the PSNR, as a classic evaluation method of FR-IQA, achieves the third rank in the overall ranking performance. What makes it perform better than most NR-IQA methods may be that it is based on the error between corresponding pixels; that is, it belongs to an error-sensitive image quality evaluation method. However, it does not take the visual properties of the human eye into account, and our algorithm still has the advantage.

Compared to the general-purpose IQA methods DIIVINE, BLIINDS, BRISQUE, CORNIA, and NIQE, the predictive values given by our algorithm in the three image databases are more consistent with the subjective quality scores. To be specific, the most obvious advantage is shown in the TID2013 database, where the general-purpose IQA methods according to the two criteria of PLCC and SRCC are only around 0.8, while those of our algorithm are higher than 0.9. Moreover, even for the SRCC metric in the LIVE database with the smallest ranking gap, our approach is at least two places ahead. Importantly, the proposed method has a better advantage in comparison with the noise-specific IQA methods PWN, DMDM, GSVD, and QENI. The overall performance of our method in the three image databases is much better than those of other methods except for the BNIQAK. Compared with the BNIQAK, our algorithm has the advantage in most PLCC values reflecting the accuracy, and our algorithm has a slight deficiency in the SRCC values reflecting the monotonicity. In most practical IQA applications, compared with SRCC, the prediction accuracy is dominant. Therefore, our proposed method has better practical values because of its higher prediction accuracy.

Table 3 Comparison of the proposed method with state-of-the-art methods on the LIVE, CSIQ, and TID2013 databases

Method	PLCC (rank)			SRCC (rank)			Average PLCC rank (rank)
	LIVE	TID2013	CSIQ	LIVE	TID2013	CSIQ	
PSNR	0.9878 (3)	0.9073 (3)	0.9363 (4)	0.9854 (2)	0.9046 (3)	0.9385 (4)	3.33 (3)
SSIM	0.9693 (9)	0.7843 (7)	0.8974 (6)	0.9694 (9)	0.7865 (8)	0.8952 (6)	7.33 (7)
DIIVINE	0.9623 (11)	0.7043 (12)	0.8118 (11)	0.9167 (12)	0.7090 (12)	0.8313 (10)	11.33 (12)
BLIINDS	0.9482 (12)	0.7296 (11)	0.8825 (7)	0.9477 (11)	0.7280 (11)	0.8863 (7)	10.00 (11)
BRISQUE	0.9780 (6)	0.7824 (8)	0.9292 (5)	0.9786 (5)	0.7784 (9)	0.9310 (5)	6.33 (5)
CORNIA	0.9870 (4)	0.7611 (10)	0.7523 (12)	0.9760 (7)	0.7561 (10)	0.7458 (12)	8.67 (9)
NIQE	0.9773 (7)	0.8255 (6)	0.8182 (10)	0.9718 (8)	0.8194 (5)	0.8098 (11)	7.67 (8)
PWN	0.9770 (8)	0.8262 (5)	0.8823 (8)	0.9816 (4)	0.8184 (6)	0.8752 (8)	7.00 (6)
DMDM	0.9789 (5)	0.8987 (4)	0.9383 (3)	0.9782 (6)	0.8953 (4)	0.9387 (3)	4.00 (4)
GSVD	0.9666 (10)	0.7731 (9)	0.8721 (9)	0.9618 (10)	0.7942 (7)	0.8480 (9)	9.33 (10)
QENI	0.8919 (13)	0.5322 (13)	0.6242 (13)	0.8828 (13)	0.3540 (13)	0.5978 (13)	13.00 (13)
BNIQAK	0.9904 (2)	0.9543 (1)	0.9550 (2)	0.9889 (1)	0.9411 (1)	0.9594 (1)	1.67 (2)
Proposed	0.9909 (1)	0.9294 (2)	0.9638 (1)	0.9830 (3)	0.9187 (2)	0.9486 (2)	1.33 (1)

Next, to intuitively evaluate the influence of the number of training images on the proposed algorithm, four groups of training images with different numbers are set in the three image databases. The ratios of the training to test images in the same image database are set to 80%/20%, 70%/30%, 60%/40%, 50%/50%, and 40%/60%. The results are listed in Table 4, and the subjective scores versus the proposed model's prediction scores on different training/testing proportions in the three databases are displayed in Fig. 9. With the reduction in the training images, the two criteria decrease slightly. Nevertheless, overall, the algorithm achieves good results. It is worth noting that even when the proportion of the training set is less than that of the testing set, the proposed algorithm still performs well in the three databases, and the prediction accuracy (PLCC) can reach higher than 0.9. This work is of significance in practical applications because it can achieve ideal prediction results even when there are insufficient training images.

Table 4 Performance of the proposed algorithm under different training/testing proportions

Training/testing proportion	Database	PLCC	SRCC
80%/20%	LIVE	**0.9909**	**0.9830**
	TID2013	**0.9294**	**0.9187**
	CSIQ	**0.9638**	0.9486
70%/30%	LIVE	0.9901	0.9802
	TID2013	0.9198	0.9045
	CSIQ	0.9616	0.9476
60%/40%	LIVE	0.9889	0.9794
	TID2013	0.9174	0.9011
	CSIQ	0.9597	0.9463
50%/50%	LIVE	0.9875	0.9774
	TID2013	0.9141	0.9002
	CSIQ	0.9592	0.9458
40%/60%	LIVE	0.9791	0.9771
	TID2013	0.9196	0.8945
	CSIQ	0.9608	**0.9492**

Best results are in bold

For the algorithm based on learning to predict the image quality score, the generalizability of the algorithm is also a significant aspect (Li LD et al., 2017). Therefore, we use cross-validation to validate the prediction performance of the proposed method, as presented in Table 5. Specifically, three different image databases are set as the training image database in turn, and the performance is tested in the other image databases. Because the level of noise distortion of images in the TID2013 and CSIQ databases is much less than that in the LIVE image database, the TID2013 and CSIQ databases are unsuitable as a training set to test the LIVE database. Table 5 indicates that all test results are greater than 0.88, the TID2013 image database exhibits the best performance as the training set, and the CSIQ has the best performance as the testing set and can achieve a PLCC of 0.9599 and an SRCC of 0.9494. The results reveal that the method can be applied to the prediction of different data sets and has good generalization ability.

Table 5 Cross-validation of the proposed method

Training database	Testing database	PLCC	SRCC
LIVE	CSIQ	0.9459	0.9458
	TID2013	0.8857	0.8847
TID2013	CSIQ	**0.9599**	**0.9494**
	LIVE	–	–
CSIQ	TID2013	0.9165	0.9109
	LIVE	–	–

Best results are in bold. "–" means that the TID2013 and CSIQ databases are unsuitable as a training set to test the LIVE database

In our method, three features are extracted for model training, and we evaluate the contribution of different features to the model by randomly combining the features. The results are listed in Table 6, where the individual results of features $\mathcal{K}$ and $\mathcal{G}$ have unsatisfactory results. However, when these two features are combined, the combination (i.e., $\mathcal{K}+\mathcal{G}$) presents encouraging results and dramatically improves the overall performance. However, the performance of a single feature or the combination of two features is not the optimal. When three features are all involved during the training procedure, the test results are the best among all experimental results. This indicates that the three extracted features are essential for the NR noisy image assessment.

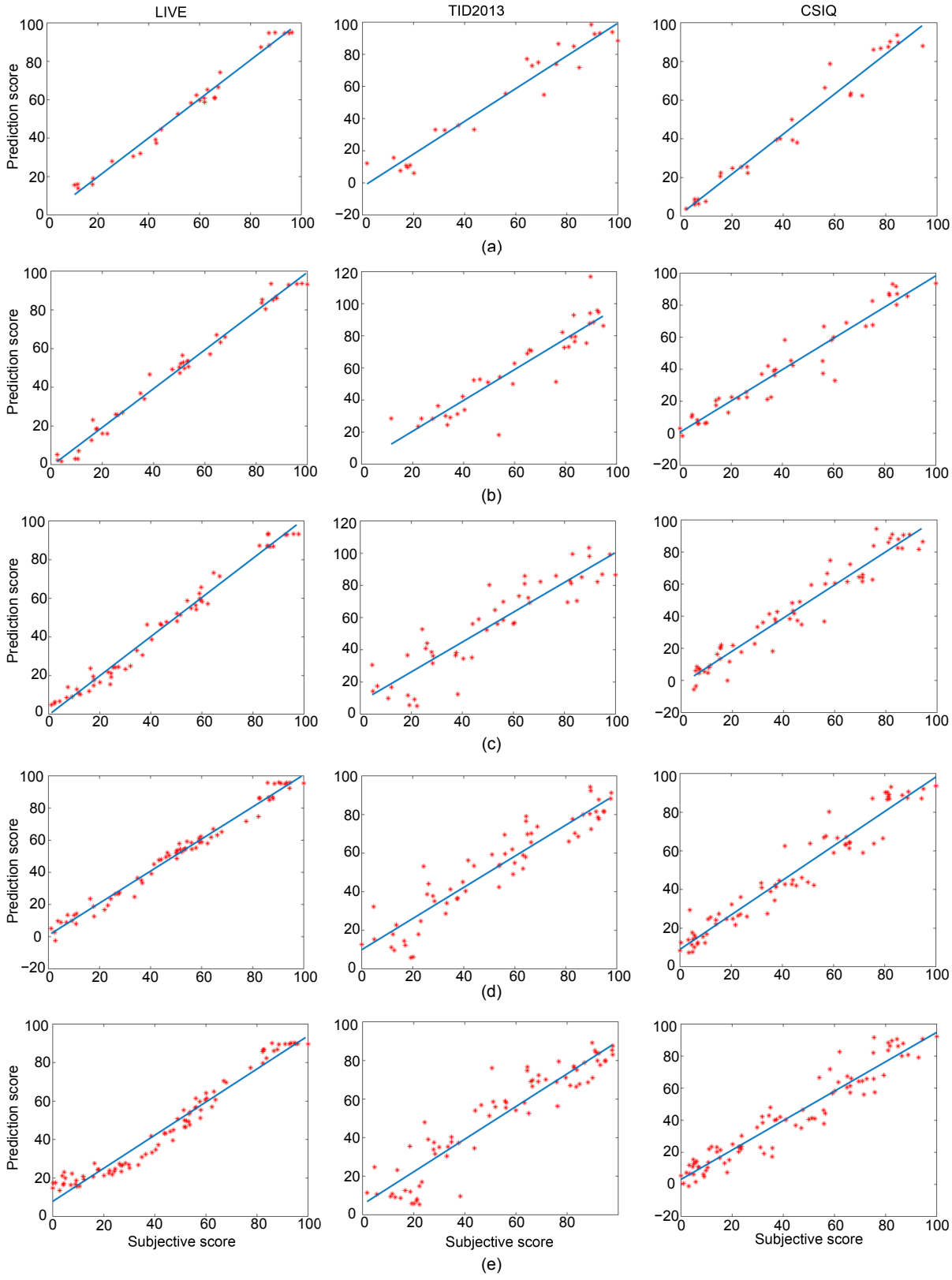


Fig. 9 Prediction score versus subjective score on different training/testing proportions in the three databases: (a) 80%/20%; (b) 70%/30%; (c) 60%/40%; (d) 50%/50%; (e) 40%/60%

Table 6 Performance comparison with different feature or feature combinations

Feature or feature combination	Database	PLCC	SRCC
$\mathcal{H}$	LIVE	0.9868	0.9764
	TID2013	0.8941	0.8875
	CSIQ	0.9223	0.8974
$\mathcal{G}$	LIVE	0.9260	0.8872
	TID2013	0.5244	0.5152
	CSIQ	0.5643	0.5433
$\mathcal{K}$	LIVE	0.9034	0.8763
	TID2013	0.5558	0.5434
	CSIQ	0.7445	0.6977
$\mathcal{H}+\mathcal{G}$	LIVE	0.9901	0.9814
	TID2013	0.9023	0.8928
	CSIQ	0.9463	0.9355
$\mathcal{K}+\mathcal{H}$	LIVE	0.9865	0.9730
	TID2013	0.8983	0.8934
	CSIQ	0.9387	0.9003
$\mathcal{K}+\mathcal{G}$	LIVE	0.9334	0.9133
	TID2013	0.7497	0.7376
	CSIQ	0.7844	0.7701
$\mathcal{K}+\mathcal{H}+\mathcal{G}$	LIVE	**0.9909**	**0.9830**
	TID2013	**0.9294**	**0.9187**
	CSIQ	**0.9638**	**0.9486**

Best results are in bold

5 Conclusions

In this paper, we designed a blind IQA metric for noisy images incorporating the features of entropy, gradient, and kurtosis. The three relevant features were extracted from different domains, and the gradient variation coefficient that can establish relationships between different domains was designed in combination with the masking effect of HVS. Finally, SVR was used to learn to map all extracted features into perceptual quality scores. Images contaminated with AWGN in the LIVE, TID2013, and CSIQ databases were used to evaluate the performance of the proposed method. We compared the FR-IQA methods, general-purpose NR-IQA methods, and noise-specific NR-IQA methods. The results demonstrated the efficacy and superiority of the proposed method, especially under the criterion of PLCC. Compared with the BNIQAK, the proposed algorithm has a slight improvement on the PLCC in the average performance. That is probably because of the comprehensive feature selection and the consideration of HVS. The evaluation results also verified the merits of our method with high precision and satisfactory generalizability, and had less influence from the number of training images.

Contributors

Heng YAO designed the algorithms. Ben MA drafted the manuscript. Mian ZOU processed the data. Dong XU and Jincao YAO supervised the algorithms, and revised and finalized the paper.

Compliance with ethics guidelines

Heng YAO, Ben MA, Mian ZOU, Dong XU, and Jincao YAO declare that they have no conflict of interest.

References

Bosse S, Maniry D, Wiegand T, et al., 2016. A deep neural network for image quality assessment. Proc IEEE Int Conf on Image Processing, p.3773-3777. https://doi.org/10.1109/ICIP.2016.7533065

Buczkowski M, 2018. Non-reference image quality assessment based on noise estimation. Proc 25th Int Conf on Systems, Signals and Image Processing, p.1-4. https://doi.org/10.1109/IWSSIP.2018.8439331

Chang CC, Lin CJ, 2011. LIBSVM: a library for support vector machines. *ACM Trans Intell Syst Technol*, 2(3):27. https://doi.org/10.1145/1961189.1961199

Chen DQ, Wang YZ, Gao W, 2020. No-reference image quality assessment: an attention driven approach. *IEEE Trans Image Process*, 29:6496-6506. https://doi.org/10.1109/TIP.2020.2990342

Deng CW, Wang SG, Bovik AC, et al., 2020. Blind noisy image quality assessment using sub-band kurtosis. *IEEE Trans Cybern*, 50(3):1146-1156. https://doi.org/10.1109/TCYB.2018.2889376

Ding Y, Li N, Zhao Y, et al., 2016. Image quality assessment method based on nonlinear feature extraction in kernel space. *Front Inform Technol Electron Eng*, 17(10):1008-1017. https://doi.org/10.1631/FITEE.1500439

Dong L, Zhou JT, Tang YY, 2017. Noise level estimation for natural images based on scale-invariant kurtosis and piecewise stationarity. *IEEE Trans Image Process*, 26(2): 1017-1030. https://doi.org/10.1109/TIP.2016.2639447

Gu K, Zhai GT, Yang XK, et al., 2015. Using free energy principle for blind image quality assessment. *IEEE Trans Multim*, 17(1):50-63. https://doi.org/10.1109/TMM.2014.2373812

Guo R, Shen XJ, Dong XY, et al., 2020. Multi-focus image fusion based on fully convolutional networks. *Front Inform Technol Electron Eng*, 21(7):1019-1033. https://doi.org/10.1631/FITEE.1900336

Hu B, Li LD, Wu JJ, et al., 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

Huang XT, Chen L, Tian J, et al., 2014. Blind noisy image quality assessment using block homogeneity. *Comput Electr Eng*, 40(3):796-807. https://doi.org/10.1016/j.compeleceng.2013.08.002

Jiang XH, Shen LQ, Yu LW, et al., 2020. No-reference screen content image quality assessment based on multi-region features. *Neurocomputing*, 386:30-41. https://doi.org/10.1016/j.neucom.2019.12.027

Kennedy J, Eberhart R, 1995. Particle swarm optimization. Proc Int Conf on Neural Networks, p.1942-1948. https://doi.org/10.1109/ICNN.1995.488968

Kong XF, Li K, Yang QX, et al., 2013. A new image quality metric for image auto-denoising. Proc IEEE Int Conf on Computer Vision, p.2888-2895. https://doi.org/10.1109/ICCV.2013.359

Larson EC, Chandler DM, 2010. Most apparent distortion: full-reference image quality assessment and the role of strategy. *J Electron Image*, 19(1):011006. https://doi.org/10.1117/1.3267105

Li LD, Xia WH, Fang YM, et al., 2016a. Color image quality assessment based on sparse representation and reconstruction residual. *J Vis Commun Image Represent*, 38: 550-560. https://doi.org/10.1016/j.jvcir.2016.04.006

Li LD, Lin WS, Wang XS, et al., 2016b. No-reference image blur assessment based on discrete orthogonal moments. *IEEE Trans Cybern*, 46(1):39-50. https://doi.org/10.1109/TCYB.2015.2392129

Li LD, Xia WH, Lin WS, et al., 2017. No-reference and robust image sharpness evaluation based on multiscale spatial and spectral features. *IEEE Trans Multim*, 19(5): 1030-1040. https://doi.org/10.1109/TMM.2016.2640762

Li PY, Lo KT, 2018. A content-adaptive joint image compression and encryption scheme. *IEEE Trans Multim*, 20(8): 1960-1972. https://doi.org/10.1109/TMM.2017.2786860

Li QH, Lin WS, Fang YM, 2017. BSD: blind image quality assessment based on structural degradation. *Neurocomputing*, 236:93-103. https://doi.org/10.1016/j.neucom.2016.09.105

Liu M, Zhai GT, Zhang ZY, et al., 2014. Blind image quality assessment for noise. Proc IEEE Int Symp on Broadband Multimedia Systems and Broadcasting, p.1-5. https://doi.org/10.1109/BMSB.2014.6873480

Lyu SW, Pan XY, Zhang X, 2014 Exposing region splicing forgeries with blind local noise estimation. *Int J Comput Vis*, 110(2):202-221. https://doi.org/10.1007/s11263-013-0688-y

Ma B, Yao JC, Le YF, et al., 2020. Efficient image noise estimation based on skewness invariance and adaptive noise injection. *IET Image Process*, 14(7):1393-1401. https://doi.org/10.1049/iet-ipr.2019.1548

Min XK, Zhai GT, Gu K, et al., 2018. Blind image quality estimation via distortion aggravation. *IEEE Trans Broadcast*, 64(2):508-517. https://doi.org/10.1109/TBC.2018.2816783

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

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

Moorthy AK, Bovik AC, 2011. Blind image quality assessment: from natural scene statistics to perceptual quality. *IEEE Trans Image Process*, 20(12):3350-3364. https://doi.org/10.1109/TIP.2011.2147325

Ospina-Borras JE, Restrepo HDB, 2016. Non-reference assessment of sharpness in blur/noise degraded images. *J Vis Commun Image Represent*, 39:142-151. https://doi.org/10.1016/j.jvcir.2016.05.015

Oszust M, 2019. No-reference quality assessment of noisy images with local features and visual saliency models. *Inform Sci*, 482:334-349. https://doi.org/10.1016/j.ins.2019.01.034

Pan CH, Xu Y, Yan YC, et al., 2016. Exploiting neural models for no-reference image quality assessment. Proc Visual Communications and Image Processing, p.1-4. https://doi.org/10.1109/VCIP.2016.7805524

Ponomarenko N, Ieremeiev O, Lukin V, et al., 2013. A new color image database TID2013: innovations and results. Proc 15th Int Conf on Advanced Concepts for Intelligent Vision Systems, p.402-413. https://doi.org/10.1007/978-3-319-02895-8_36

Saad MA, Bovik AC, Charrier C, 2012. Blind image quality assessment: a natural scene statistics approach in the DCT domain. *IEEE Trans Image Process*, 21(8):3339-3352. https://doi.org/10.1109/TIP.2012.2191563

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

Shen LL, Hang N, Hou CP, 2020. Feature-segmentation strategy based convolutional neural network for no-reference image quality assessment. *Multim Tool Appl*, 79(17-18):11891-11904. https://doi.org/10.1007/s11042-019-08298-2

Tang LJ, Li LD, Sun KZ, et al., 2017. An efficient and effective blind camera image quality metric via modeling quaternion wavelet coefficients. *J Vis Commun Image Represent*, 49:204-212. https://doi.org/10.1016/j.jvcir.2017.09.010

Tang ZJ, Huang ZQ, Yao H, et al., 2018. Perceptual image hashing with weighted DWT features for reduced-reference image quality assessment. *Comput J*, 61(11): 1695-1709. https://doi.org/10.1093/comjnl/bxy047

Video Quality Experts Group, 2003. Final Report from the Video Quality Experts Group on the Validation of Objective Models of Video Quality Assessment, Phase II (fr_tv2). http://www.vqeg.org

Wang Q, Chu J, Xu L, et al., 2016. A new blind image quality framework based on natural color statistic. *Neurocomputing*, 173:1798-1810.

https://doi.org/10.1016/j.neucom.2015.09.057

Wang Z, Bovik AC, Sheikh HR, et al., 2004. Image quality assessment: from error visibility to structural similarity. *IEEE Trans Image Process*, 13(4):600-612. https://doi.org/10.1109/TIP.2003.819861

Wu JJ, Zhang M, Li LD, et al., 2019. No-reference image quality assessment with visual pattern degradation. *Inform Sci*, 504:487-500. https://doi.org/10.1016/j.ins.2019.07.061

Xu L, Huang G, Chen QL, et al., 2020. An improved method for image denoising based on fractional-order integration. *Front Inform Technol Electron Eng*, 21(10):1485-1493. https://doi.org/10.1631/FITEE.1900727

Ye P, Kumar J, Kang L, et al., 2012. Unsupervised feature learning framework for no-reference image quality assessment. Proc IEEE Conf on Computer Vision and Pattern Recognition, p.1098-1105. https://doi.org/10.1109/CVPR.2012.6247789

Zhai GT, Wu XL, 2011. Noise estimation using statistics of natural images. Proc 18th IEEE Int Conf on Image Processing, p.1857-1860. https://doi.org/10.1109/ICIP.2011.6115828

Zhai GT, Wu XL, Yang XK, et al., 2012. A psychovisual quality metric in free-energy principle. *IEEE Trans Image Process*, 21(1):41-52. https://doi.org/10.1109/TIP.2011.2161092

Zhai GT, Kaup A, Wang J, et al., 2015. A dual-model approach to blind quality assessment of noisy images. *APSIPA Trans Signal Inform Process*, 4:e4. https://doi.org/10.1017/ATSIP.2015.8

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

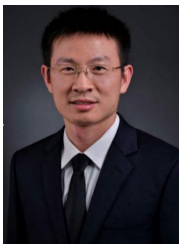
Zhou WJ, Yu L, Qiu WW, et al., 2017. Local gradient patterns (LGP): an effective local-statistical-feature extraction scheme for no-reference image quality assessment. *Inform Sci*, 397-398:1-14. https://doi.org/10.1016/j.ins.2017.02.049

Zhu HC, Li LD, Wu JJ, et al., 2020. MetaIQA: deep meta-learning for no-reference image quality assessment. IEEE/CVF Conf on Computer Vision and Pattern Recognition, p.14143-14152. https://doi.org/10.1109/CVPR42600.2020.01415

Zhu T, Karam L, 2014. A no-reference objective image quality metric based on perceptually weighted local noise. *EURASIP J Image Video Process*, 2014(1):1-8. https://doi.org/10.1186/1687-5281-2014-5

Zoran D, Weiss Y, 2009. Scale invariance and noise in natural images. Proc IEEE Int Conf on Computer Vision, p.2209-2216. https://doi.org/10.1109/ICCV.2009.5459476

Heng YAO, first author of this invited paper, received his BS degree from Hefei University of Technology, China, in 2004, his MS degree from Shanghai Normal University, China, in 2008, and his PhD degree in Signal and Information Processing from Shanghai University, China, in 2012. He is currently an associate professor at University of Shanghai for Science and Technology. His research interests include multimedia security, image processing, and pattern recognition. He has published more than 40 peer-reviewed papers in international journals.



Dong XU, corresponding author of this invited paper, received his BS and MS degrees in Medical Imaging from Southeast University and Zhejiang University in 2002 and 2008, respectively. In 2011, he received his PhD degree in Clinical Specialty of Integrated Traditional Chinese and Western Medicine from Zhejiang Chinese Medical University. He is now a professor at the Institute of Basic Medicine and Cancer (IBMC), Chinese Academy of Sciences and Cancer Hospital of the University of Chinese Academy of Sciences (Zhejiang Cancer Hospital). His current research interests include medical image processing, machine learning, radiomics, and medical imaging diagnosis of tumor.

Jincao YAO, corresponding author of this invited paper, received his BS degree in Computer and Information Science from Hefei University of Technology in 2004, and his MS degree in Signal Processing from Shanghai Normal University in 2007. In 2017, he received his PhD degree in Signal and Information Processing at the College of Information Science and Electronic Engineering, Zhejiang University, China. He is now an imaging physicist at the IBMC, Chinese Academy of Sciences and Cancer Hospital of the University of Chinese Academy of Sciences (Zhejiang Cancer Hospital). His current research interests include machine learning, computer vision, medical image recognition, radiomics, and shape-driven techniques in image processing.