

# UNSUPERVISED BLIND IMAGE QUALITY ASSESSMENT BASED ON MULTI-FEATURE FUSION

Qinglin He, Chao Yang and Ping An

School of Communication and Information Engineering,  
Shanghai University, Shanghai, China

## **ABSTRACT**

*Image quality affects the visual experience of observers. How to accurately evaluate image quality has been widely studied by researchers. Unsupervised blind image quality assessment (BIQA) requires less prior knowledge than supervised ones. Besides, there is a trade-off between accuracy and complexity in most existing BIQA methods. In this paper, we propose an unsupervised BIQA framework that aims for both high accuracy and low complexity. To represent the image structure information, we employ Phase Congruency (PC) and gradient. After that, we calculate the mean subtracted and contrast normalized (MSCN) coefficient and the Karhunen-Loève transform (KLT) coefficient to represent the naturalness of the images. Finally, features extracted from both the pristine and the distorted images are adopted to calculate the image quality with Multivariate Gaussian (MVG) model. Experiments conducted on six IQA databases demonstrate that the proposed method achieves better performance than the state-of-the-art BIQA methods.*

## **KEYWORDS**

*Blind Image Quality Assessment (BIQA), Unsupervised Method, Natural Scene Statistics (NSS), Karhunen-Loève Transform (KLT).*

## **1. INTRODUCTION**

With the rapid development of multimedia, the quality of images not only affects the visual experience of observers but also has an impact on image processing algorithms. How to measure image quality with lower computational complexity and better generalization performance has been a hot spot. The score of each image is the main evaluation criteria. There are two primary categories of score acquiring, namely the subjective evaluation and the objective evaluation. For the subjective evaluation, scores of different people on the same picture are needed and that is costly and time-consuming. On the other hand, models without human involvement are easy to use on large-scale databases for objective evaluation. The goal of image quality assessment (IQA) is to fit the objective score as close as possible to the subjective score, which means we could extract well-chosen features to imitate human behavior for more precise scores.

In general, the objective image quality assessment can be classified into three types which are full-reference (FR) IQA [1-6], reduced-reference (RR) IQA [7-10], and no-reference (NR) IQA [11-14]. FR IQA methods need the original image and its distorted version to fit the model, and RR IQA methods need features of the original image. NR IQA, which is also called Blind IQA (BIQA), only needs distorted images to predict scores. In FR and RR IQA methods, the need for

the original image as a reference limits the practical use. On the contrary, the BIQA methods do not have such strict requirements for model fitting and evaluation.

BIQA can be divided into supervised and unsupervised approaches. Supervised approaches usually utilize subjective scores as the ground truth to train the model. Mittal *et al.* [11] extracted natural scene statistics (NSS) features from local normalized images and Yang *et al.* [15] employed Karhunen-Loève transform (KLT) for learning-based features extraction, then these features were projected to subjective scores using support vector regression. Zeng *et al.* [16] used probabilistic quality representation and a more robust function for training the deep BIQA model. Ma *et al.* [17] proposed a multi-task learning-based deep learning approach, which consists of distortion identification and quality prediction tasks. Zhu *et al.* [18] proposed a deep meta-learning model for prior knowledge learning with good generalization ability. By simulating the human visual system, Chang *et al.* [19] used a visual neuron matrix (VNM) evaluator for quality assessment.

Unlike supervised approaches, unsupervised approaches can reveal better generalization capability with few manual calibration data. Wu *et al.* [13] proposed a highly efficient method for real-time evaluation. Wu *et al.* [12] proposed a visual perception nature image quality evaluation model for score training, which had an understanding-based global-local structure to simulate the top-down structure. Natural image quality evaluator (NIQE) [14] and its feature enriched extension, integrated local NIQE (ILNIQE) [20] introduced multivariate Gaussian (MVG) model for BIQA which required no subjective scores for regression model training. Liu *et al.* [21] introduced structure, naturalness, and perception features to the NIQE framework for further study.

In this paper, we propose a multi-feature fusion NIQE with better performance and lower complexity. We select Phase Congruency (PC) and gradient as structure features and select mean subtracted and contrast normalized (MSCN) and KLT coefficient as Natural Scene Statistics features. Fused structure features and NSS features are used for the MVG model fitting. Experiments show that the proposed unsupervised method achieves better performance with lower computational complexity on different databases. The rest of this paper is organized as follows. Section II introduces the detailed framework of our method. Section III reports the experimental results, and Section IV concludes this paper.

## 2. PROPOSED METHOD

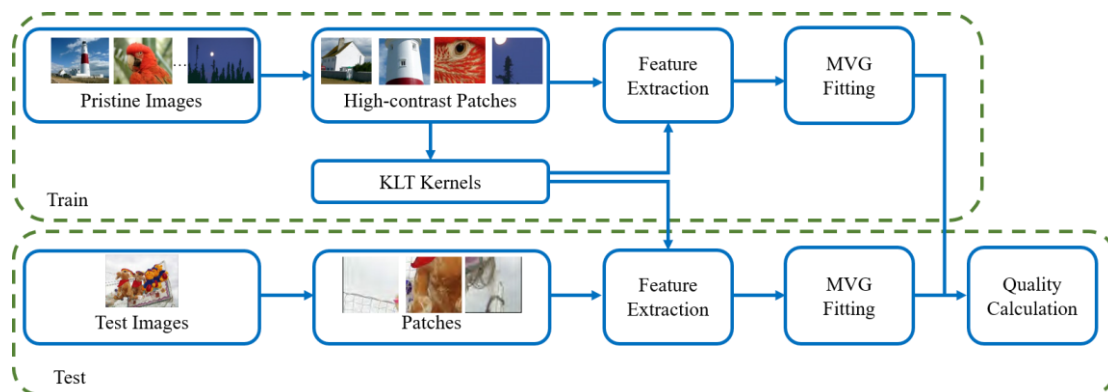


Figure 1. Framework of the proposed method

The framework of the proposed method is shown in Figure 1. To form the feature matrix, we extract structure features and NSS features from each non-overlapping patch of the image, which contains PC and gradient in the former, MSCN together with KLT coefficient in the latter. Then, we fit the MVG model with the feature matrix of pristine images as a benchmark. The distance between the benchmark model and the MVG model of the distorted image is taken as the objective score.

## 2.1. Structure features

### 2.1.1. Phase Congruency

Phase congruency calculates the maximum moment of PC covariance, which is used as an indicator of edge strength. We utilize [22] to compute the PC map of an image. For 2D signal  $s$ , the responses of even and odd-symmetric filters at position  $p$  can be denoted as

$$[e_{n,\theta_f(p)}, o_{n,\theta_f(p)}], \quad (1)$$

where  $n$  and  $\theta_f$  refer to scale and direction respectively, and  $\theta_f = f\pi / F$ ,  $f = 0, 1, \dots, F-1$  where  $F$  is the number of filter directions. The local amplitude is  $A_{n,\theta_f}(p) = \sqrt{e_{n,\theta_f}(p)^2 + o_{n,\theta_f}(p)^2}$ . Let  $E_{n,\theta_f}(p) = \sum_n e_{n,\theta_f}(p)$ ,  $O_{n,\theta_f}(p) = \sum_n o_{n,\theta_f}(p)$ . Phase congruency is calculated using:

$$PC_{2D}(p) = \frac{\sum_f H_{\theta_f}(p)}{\varepsilon + \sum_f \sum_n A_{n,\theta_f}(p)} \quad (2)$$

where  $H_{\theta_f}(p) = \sqrt{E_{n,\theta_f}(p)^2 + O_{n,\theta_f}(p)^2}$  and  $\varepsilon$  is a small positive constant.

We calculate the PC feature from the color relevant space  $O$ , which is converted from RGB in [23]:

$$\begin{bmatrix} O_1 \\ O_2 \\ O_3 \end{bmatrix} = \begin{bmatrix} 0.06 & 0.63 & 0.27 \\ 0.30 & 0.04 & -0.35 \\ 0.34 & -0.60 & 0.17 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (3)$$

Finally, Weibull distribution is used to model the PC distribution in each color channel with scale parameter  $\lambda$  and shape parameter  $q$ . The dimension of the feature is  $1 \times 6$ :

$$f(x; \lambda, q) = \begin{cases} \frac{q}{\lambda} \left(\frac{x}{\lambda}\right)^{q-1} \exp\left(-\left(\frac{x}{\lambda}\right)^q\right) & x \geq 0 \\ 0 & x < 0 \end{cases} \quad (4)$$

### 2.1.2. Image Gradient

The gradient is an indispensable IQA index that represents the contrast and luminance information of an image. We use filters  $D_h = [1, -1]$  and  $D_v = [1, -1]^T$  to compute the horizontal and vertical gradient:

$$\begin{cases} G_h = I * D_h \\ G_v = I * D_v \end{cases} \quad (5)$$

where  $I$  refers to the image patch in the luminance channel and  $*$  is the convolution operation. The distribution of  $G_h$  and  $G_v$  could be modeled as zero-mean General Gaussian Distribution (GGD) in [21]:

$$f(x; \alpha, \sigma^2) = \frac{\alpha}{2\beta(1/\alpha)} \exp\left(-\left(\frac{|x|}{\beta}\right)^\alpha\right) \quad (6)$$

where  $\beta$  refers to standard deviation  $\beta = \sigma \sqrt{\frac{\Gamma(1/\alpha)}{\Gamma(3/\alpha)}}$ , and  $\Gamma(\cdot)$  refers to gamma function, defined as:

$$\Gamma(a) = \int_0^\infty t^{a-1} e^{-t} dt \quad a > 0 \quad (7)$$

We use  $\alpha$  and  $\sigma$  as features and get a feature vector with a dimension of  $1 \times 4$ .

## 2.2. NSS features

### 2.2.1. Mean Subtracted and Contrast Normalized Coefficient

Natural images and distorted images have different MSCN coefficient distributions. The extraction of MSCN coefficients from the image patch  $I$  in the luminance channel is as follows:

$$M(i, j) = \frac{I(i, j) - \mu(i, j)}{\sigma(i, j) + 1} \quad (8)$$

$$\mu(i, j) = \sum_{l=-3}^3 \sum_{k=-3}^3 \omega_{l,k} I(i+k, j+l) \quad (9)$$

$$\sigma^2(i, j) = \sum_{l=-3}^3 \sum_{k=-3}^3 \omega_{l,k} [I(i+k, j+l) - \mu(i, j)]^2 \quad (10)$$

where  $i$  and  $j$  indicate the coordinates of the pixel and  $\{\omega = \omega_{l,k} \mid l = -3, \dots, 3, k = -3, \dots, 3\}$  defines a unit-volume Gaussian window.

GGD in Eq. (6) is used to fit MSCN distribution and get a 2-dimension feature. Furthermore, the asymmetric generalized Gaussian distribution (AGGD) model is applied to fit adjacent MSCN coefficients along with four directions, *i.e.* horizontal, vertical, main diagonal, and sub-diagonal [11]. Wherein the AGGD model is calculated:

$$f(x; \gamma, \beta_l, \beta_r) = \begin{cases} \frac{\gamma}{(\beta_l + \beta_r)\Gamma(1/\gamma)} \exp(-(\frac{-x}{\beta_l})^\gamma) \forall x \leq 0 \\ \frac{\gamma}{(\beta_l + \beta_r)\Gamma(1/\gamma)} \exp(-(\frac{-x}{\beta_r})^\gamma) \forall x \geq 0 \end{cases} \quad (11)$$

where  $\gamma$  controls the shape,  $\beta_l$  and  $\beta_r$  represent left and right side scale respectively. The mean value of this distribution is calculated as below:

$$\eta = (\beta_r - \beta_l) \frac{\Gamma(2/\gamma)}{\Gamma(1/\gamma)} \quad (12)$$

The parameters  $(\gamma, \beta_l, \beta_r, \eta)$  for adjacent MSCN coefficient are set as features with a dimension of  $1 \times 16$ .

### 2.2.2. Karhunen-Loève Transform

KLT is a data-driven feature extractor to extract image structural features [15]. Non-overlapping patches of MSCN normalized pristine image  $M$  with size  $\sqrt{k} \times \sqrt{k}$  is used to collect vectorized patch  $r_u$ , wherein  $u$  is the patch index, and the covariance matrix  $C$  is defined as:

$$\begin{aligned} C &= E[(r - m)(r - m)^T] \\ &= \frac{1}{U} \sum_{u=1}^U [(r_u - m)(r_u - m)^T], \end{aligned} \quad (13)$$

where  $U$  is the number of training patches, and  $m$  is the average vector of each vectorized patch. Then, the KLT kernel with size  $k \times k$  is the eigenvectors of the covariance matrix in Eq. (13), denoted as  $P$ . Each column vector in  $P$  is an eigenvector of the covariance matrix  $C$ , and these eigenvectors are arranged in descending order according to their eigenvalues.

GGD is chosen to fit the KLT coefficient distribution, the kernel size  $k$  is set to 4 and the feature dimension is  $1 \times 8$ .

### 2.3. MVG fitting for Unsupervised BIQA

The features extracted above can be fitted with the MVG model as follows:

$$f(x) = \frac{1}{(2\pi)^{d/2} |\Sigma|^{1/2}} \exp(-\frac{1}{2} (x - \nu)^T \Sigma^{-1} (x - \nu)) \quad (14)$$

where  $x$  represents the feature extracted from the image patches and  $d$  is the dimension of the feature vector,  $\nu$  and  $\Sigma$  refer to the mean vector and covariance matrix of  $x$  respectively. In this paper, all features are extracted at two scales, *i.e.* the original image scale, and the down-sampled scale by a factor of 2. The corresponding image patch size of  $I$  is  $96 \times 96$  and  $48 \times 48$ . Therefore, the dimension of the features extracted from each image patch is  $1 \times 72$ .

Then, the quality of the distorted image is measured as the distance between MVG parameters of the pristine images and distorted image:

$$Q = \sqrt{(\nu_1 - \nu_2)^T \left( \frac{\Sigma_1 + \Sigma_2}{2} \right)^{-1} (\nu_1 - \nu_2)} \quad (15)$$

where  $\nu_1$ ,  $\nu_2$  and  $\Sigma_1$ ,  $\Sigma_2$  are the mean vectors and covariance matrices of the pristine MVG model and the distorted image's MVG model.

### 3. EXPERIMENTS

#### 3.1. Databases and Evaluation Methodology

Six widely utilized IQA databases including LIVE [25], MICT [26], CSIQ [27], TID2013 [28], CID2013 [29] and LIVE Challenge [30] *i.e.* LIVE-C are used to test the performance of the proposed method. We utilize the full LIVE and MICT databases for experiments. While for CSIQ and TID2013, we test on common distortion types, *i.e.* JPEG, JPEG2000, White Noise, and Gaussian Blur for a fair comparison. LIVE-C and CID2013 have real-world distortion without specific distortion types, therefore we test on the whole database respectively. We employ 125 images in [14] to train the KLT kernels and fit the pristine MVG model.

#### 3.2. Overall Performance on 6 Databases

For the supervised models, we use the full LIVE database to train and then test the model on the rest five databases. For a fair comparison, we choose the three most commonly used criteria for model evaluation, which are Spearman Rank Order Correlation Coefficient (SROCC), Pearson Linear Correlation Coefficient (PLCC), and Root Mean Squared Error (RMSE). We calculate the SROCC with predicted scores and subjective scores, while for the calculation of PLCC and RMSE, we mapped the objective scores to the space of subjective scores with the nonlinear mapping method in [31].

The results of unsupervised methods on LIVE are in Table 1. The proposed method reaches the best results of three criteria. Table 2 shows the SROCC of the proposed method as well as other BIQA methods. "W. A." refers to the weighted average performance over the five databases and the weights are the number of images selected in each database. The best performances of supervised and unsupervised methods are highlighted in bold. The generalization of RankIQA is pretty good among supervised models, while the weighted average performance of the proposed method is the highest among the unsupervised methods, even higher than RankIQA.

Table 1. The performance of unsupervised BIQA models on LIVE, which contains SROCC, PLCC, and RMSE.

method	SROCC	PLCC	RMSE
LPSI[13]	0.8181	0.8280	15.3184
NIQE[14]	0.9080	0.9064	11.5429
ILNIQE[20]	0.8972	0.9021	11.7913
SNP-NIQE[21]	0.9086	0.9073	11.4893
Proposed	<b>0.9121</b>	<b>0.9095</b>	<b>11.3603</b>

Table 2. SROCC results on Different Databases.

SROCC	MICT	CSIQ	TID2013	CID2013	LIVE-C	W.A.
BRISQUE[11]	0.8526	0.8842	0.8401	0.5485	0.3026	0.5866
MEON[17]	0.8919	<b>0.9300</b>	<b>0.9012</b>	0.3813	0.3640	0.6062
RankIQA[24]	<b>0.9109</b>	0.8337	0.8670	<b>0.7040</b>	<b>0.3879</b>	<b>0.6437</b>
LPSI[13]	<b>0.9005</b>	0.7711	0.7046	0.3230	0.0834	0.4180
NIQE[14]	0.8472	0.8711	0.7966	0.6568	0.4498	0.6528
ILNIQE[20]	0.7384	0.8794	0.8422	0.3057	0.4389	0.5946
SNP-NIQE[21]	0.8908	0.9024	0.8571	0.7155	0.4652	0.6879
Proposed	0.8745	<b>0.9027</b>	<b>0.8764</b>	<b>0.7753</b>	<b>0.5036</b>	<b>0.7155</b>

The SROCC results of 24-distortion-types on TID2013 are tabulated in Table 3 and the best results of each type are highlighted in bold. ‘‘Avg.’’ refers to the average score over the 24 distortions. Distortion types vary in the TID2013 database, so reaching the highest score on each type of distortion is a great challenge for models. LPSI, SNP-NIQE, and the proposed method all have six results in bold. Among them, the proposed method has the highest average score. Besides, the results of the proposed method are competitive on TID2013 for both common and uncommon distortion types.

Table 3. SROCC results on TID2013 in different distortions.

TID2013	1	2	3	4	5	6	7	8	9	10	11	12
LPSI	0.769	0.495	0.696	0.046	<b>0.925</b>	0.432	0.853	0.840	0.248	<b>0.912</b>	0.898	0.091
	0	5	8	2	<b>0</b>	4	7	8	7	<b>3</b>	8	1
NIQE	0.814	0.590	0.541	0.721	0.851	0.744	0.860	0.809	0.577	0.859	0.866	0.121
	8	6	1	1	0	7	8	7	6	7	0	6
ILNIQE	0.875	<b>0.814</b>	<b>0.923</b>	0.511	0.869	0.753	0.873	0.814	<b>0.748</b>	0.834	0.860	0.274
	9	<b>5</b>	<b>4</b>	6	1	2	2	3	<b>3</b>	6	6	3
SNP-NIQE	<b>0.885</b>	0.733	0.649	<b>0.740</b>	0.873	<b>0.799</b>	0.857	<b>0.863</b>	0.612	0.879	0.877	0.281
	<b>6</b>	0	5	<b>0</b>	0	<b>7</b>	3	<b>8</b>	8	1	6	7
Proposed	0.832	0.729	0.825	0.717	0.857	0.785	<b>0.910</b>	0.831	0.687	0.900	<b>0.910</b>	<b>0.332</b>
	7	1	7	8	5	8	<b>1</b>	1	2	7	<b>9</b>	<b>8</b>
13	14	15	16	17	18	19	20	21	22	23	24	Avg.
<b>0.6106</b>	0.052	0.137	<b>0.340</b>	0.199	0.301	0.695	0.018	0.235	<b>0.899</b>	<b>0.695</b>	0.862	0.510
	0	2	<b>9</b>	2	8	9	1	6	<b>8</b>	<b>3</b>	0	8
0.3997	0.033	0.174	0.147	0.106	<b>0.302</b>	0.678	0.053	0.754	0.760	0.568	0.851	0.549
	1	1	0	3	<b>1</b>	8	5	2	3	3	7	5
0.5228	<b>0.079</b>	0.129	0.181	0.014	0.165	0.690	<b>0.353</b>	0.828	0.748	0.680	0.864	0.600
	<b>6</b>	7	9	2	9	1	<b>8</b>	7	8	0	9	6
0.5917	0.014	0.032	0.099	0.156	0.106	<b>0.740</b>	0.208	<b>0.830</b>	0.790	0.634	0.828	0.586
	9	1	9	2	0	<b>1</b>	3	<b>0</b>	0	7	7	9
0.4267	0.000	<b>0.198</b>	0.068	<b>0.341</b>	0.255	0.681	0.299	0.826	0.794	0.636	<b>0.886</b>	<b>0.614</b>
	7	<b>8</b>	1	<b>8</b>	5	4	3	2	1	8	<b>3</b>	<b>1</b>

### 3.3. Ablation Test

To demonstrate the effectiveness of the structure and NSS features, we report the ablation test in Table 4. NSS features take the leading role while structure features play as a supplement. The combination of these two types of features can significantly improve the performance, both of these features are indispensable.

Table 4. Performance contribution of each type of feature and their combination of SROCC.

Database	structure features	NSS features	Proposed
LIVE	0.7137	0.9068	<b>0.9121</b>
MICT	0.6272	0.8713	<b>0.8745</b>
CSIQ	0.6100	0.8976	<b>0.9027</b>
TID2013	0.6063	0.8681	<b>0.8764</b>
CID2013	0.7106	0.7162	<b>0.7753</b>
LIVE-C	0.4420	0.4865	<b>0.5036</b>

### 3.4. Significance Test

To verify the statistical significance of the results, we applied t-test [25] on the prediction residuals of different objective methods.

Table 5. Statistical significance results between SROCC values. 1, 0, or -1 implies proposed method is statistical superior, comparative, or inferior to the algorithm with 95% confidence.

SROCC	LIVE	MICT	CSIQ	TID2013	CID2013	LIVE-C
BRISQUE	-	-1	-1	0	-1	-1
MEON	-	0	-1	0	1	1
RankIQA	-	0	0	0	1	0
LPSI	1	0	1	1	1	1
NIQE	0	0	0	0	1	0
ILNIQE	1	1	0	1	1	0
SNP-NIQE	0	0	0	1	1	0

'1', '0' and '-1' in Table 5 indicate that the proposed method is statistically superior, comparative, or inferior to the competing method on each database with 95% confidence. The unsupervised method is a little inadequate compared with supervised ones. The proposed method is no worse than other unsupervised methods. The proposed method has comparable performance with NIQE and SNP-NIQE, however, it is better than ILNIQE and LPSI on more than half of the databases.

### 3.5. Computation Complexity Comparison

Table 6 shows the average running time of different unsupervised BIQA methods on LIVE. These five methods are implemented on the MATLAB platform and tested on our PC with the following configuration, CPU: Intel Core i7-3770 3.40GHz Dual-Core, RAM: 8GB, and Windows system. All images in the LIVE database are utilized for the running time test. The generalization performance is good for ILNIQE, but it has higher computational complexity. The average running time of LPSI is very short, but the accuracy and generalization ability are limited. The proposed method has the highest results of six databases on a weighted average, the generalization performance and running time are competitive.

Table 6. Average running time of different unsupervised BIQA methods on LIVE.

	LPSI	NIQE	SNP-NIQE	ILNIQE	Proposed
Times(s)	0.02	0.24	5.06	5.49	1.21



## 4. CONCLUSION

In this paper, we propose an unsupervised BIQA method based on multi-feature fusion using structure features and NSS features. We extract PC, gradient, MSCN, and KLT features from non-overlapping image patches to fit the MVG feature matrix. The distance between the pristine and distorted MVG feature matrices is used as the objective score. Experiments on six IQA databases show that the proposed method achieves better performance with lower computation complexity on both common distortion types and real-world distortion. In the future, we can extend our work to uncommon distortion types.

## ACKNOWLEDGEMENTS

This work was supported in part by the NSFC under Grant 61901252.

## REFERENCES

- [1] Zhou Wang, A. C. Bovik, H. R. Sheikh and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," in *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600-612, April 2004, doi: 10.1109/TIP.2003.819861.
- [2] Z. Wang, E. P. Simoncelli and A. C. Bovik, "Multiscale structural similarity for image quality assessment," *The Thrity-Seventh Asilomar Conference on Signals, Systems & Computers*, 2003, 2003, pp. 1398-1402 Vol.2, doi: 10.1109/ACSSC.2003.1292216.
- [3] L. Zhang, L. Zhang, X. Mou and D. Zhang, "FSIM: A Feature Similarity Index for Image Quality Assessment," in *IEEE Transactions on Image Processing*, vol. 20, no. 8, pp. 2378-2386, Aug. 2011, doi: 10.1109/TIP.2011.2109730.
- [4] J. Farah, M. Hojeij, J. Chrabieh and F. Dufaux, "Full-reference and reduced-reference quality metrics based on SIFT," *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2014, pp. 161-165, doi: 10.1109/ICASSP.2014.6853578.
- [5] X. Ma, X. Jiang and X. Guo, "Full-reference image quality assessment based on the analysis of distortion process," *2017 4th International Conference on Systems and Informatics (ICSAI)*, 2017, pp. 1256-1260, doi: 10.1109/ICSAI.2017.8248471.
- [6] J. Kim and S. Lee, "Deep blind image quality assessment by employing FR-IQA," *2017 IEEE International Conference on Image Processing (ICIP)*, 2017, pp. 3180-3184, doi: 10.1109/ICIP.2017.8296869.
- [7] R. Soundararajan and A. C. Bovik, "RRED indices: Reduced reference entropic differencing framework for image quality assessment," *2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2011, pp. 1149-1152, doi: 10.1109/ICASSP.2011.5946612.
- [8] J. Wu, Y. Liu, G. Shi and W. Lin, "Saliency change based reduced reference image quality assessment," *2017 IEEE Visual Communications and Image Processing (VCIP)*, 2017, pp. 1-4, doi: 10.1109/VCIP.2017.8305041.
- [9] Y. Liu, G. Zhai, K. Gu, X. Liu, D. Zhao and W. Gao, "Reduced-Reference Image Quality Assessment in Free-Energy Principle and Sparse Representation," in *IEEE Transactions on Multimedia*, vol. 20, no. 2, pp. 379-391, Feb. 2018, doi: 10.1109/TMM.2017.2729020.
- [10] Q. Hu, Y. Sheng, L. Yang, Q. Li and L. Chai, "Reduced-Reference Image Quality Assessment for Single-Image Super-Resolution Based on Wavelet Domain," *2019 Chinese Control and Decision Conference (CCDC)*, 2019, pp. 2067-2071, doi: 10.1109/CCDC.2019.8833247.
- [11] A. Mittal, A. K. Moorthy and A. C. Bovik, "No-Reference Image Quality Assessment in the Spatial Domain," in *IEEE Transactions on Image Processing*, vol. 21, no. 12, pp. 4695-4708, Dec. 2012, doi: 10.1109/TIP.2012.2214050.
- [12] L. Wu, X. Zhang, H. Chen, D. Wang, and J. Deng, "Vp-niqe: An opinion-unaware visual perception natural image quality evaluator," *Neurocomputing*, vol. 463, pp. 17-28, 2021.
- [13] Q. Wu, Z. Wang, and H. Li, "A highly efficient method for blind image quality assessment," in *2015 IEEE International Conference on Image Processing (ICIP)*, 2015, pp. 339-343.
- [14] A. Mittal, R. Soundararajan, and A. C. Bovik, "Making a "Completely Blind" Image Quality Analyzer," *IEEE Signal Processing Letters*, vol. 20, no. 3, pp. 209-212, 2013.

- [15] C. Yang, X. Zhang, P. An, L. Shen, and C. J. Kuo, "Blind image quality assessment based on multi-scale klt," *IEEE Transactions on Multimedia*, pp. 1–1, 2020.Z.
- [16] H. Zeng, L. Zhang, and A. C. Bovik, "Blind image quality assessment with a probabilistic quality representation," in *2018 25th IEEE International Conference on Image Processing (ICIP)*, 2018, pp. 609–61
- [17] K. Ma, W. Liu, K. Zhang, Z. Duanmu, Z. Wang, and W. Zuo, "End-to-end blind image quality assessment using deep neural networks," *IEEE Transactions on Image Processing*, vol. 27, no. 3, pp. 1202–1213, 2018.
- [18] H. Zhu, L. Li, J. Wu, W. Dong, and G. Shi, "Metaiq: Deep meta-learning for no-reference image quality assessment," in *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020, pp. 14 131–14 140.
- [19] H.-W. Chang, X.-D. Bi, and C. Kai, "Blind image quality assessment by visual neuron matrix," *IEEE Signal Processing Letters*, vol. 28, pp. 1803–1807, 2021.
- [20] L. Zhang, L. Zhang, and A. C. Bovik, "A feature-enriched completely blind image quality evaluator," *IEEE Transactions on Image Processing*, vol. 24, no. 8, pp. 2579–2591, 2015.
- [21] Y. Liu, K. Gu, Y. Zhang, X. Li, G. Zhai, D. Zhao, and W. Gao, "Unsupervised blind image quality evaluation via statistical measurements of structure, naturalness, and perception," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 30, no. 4, pp. 929–943, 2020.
- [22] P. Kovesi, "Image features from phase congruency," *Videre J. Comput. Vision Res.*, vol. 1, 01 1999.
- [23] J. Geusebroek, R. van den Boomgaard, A. W. M. Smeulders, and H. Geerts, "Color invariance," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 12, pp. 1338–1350, 2001.
- [24] X. Liu, J. Van De Weijer, and A. D. Bagdanov, "Rankiqa: Learning from rankings for no-reference image quality assessment," in *2017 IEEE International Conference on Computer Vision (ICCV)*, 2017, pp. 1040–1049.
- [25] H. R. Sheikh, M. F. Sabir, and A. C. Bovik, "A statistical evaluation of recent full reference image quality assessment algorithms," *IEEE Transactions on Image Processing*, vol. 15, no. 11, pp. 3440–3451, 2006.
- [26] Y. Horita, K. Shibata, and Y. Kawayoka, "Toyama image quality evaluation database," 2011, <http://mict.eng.u-toyama.ac.jp/mictdb.html>.
- [27] E. C. Larson and D. M. Chandler, "Most apparent distortion: full-reference image quality assessment and the role of strategy," *Journal of Electronic Imaging*, vol. 19, no. 1, p. 011006, 2010.
- [28] N. Ponomarenko, O. Ieremeiev, V. Lukin, K. Egiazarian, L. Jin, J. Astola, B. Vozel, K. Chehdi, M. Carli, F. Battisti, and C. J. Kuo, "Color image database tid2013: Peculiarities and preliminary results," in *European Workshop on Visual Information Processing (EUVIP)*, 2013, pp. 106–111.
- [29] T. Virtanen, M. Nuutinen, M. Vaahteranoksa, P. Oittinen, and J. Häkkinen, "CID2013: A Database for Evaluating No-Reference Image Quality Assessment Algorithms," *IEEE Transactions on Image Processing*, vol. 24, no. 1, pp. 390–402, 2015
- [30] D. Ghadiyaram and A. C. Bovik, "Massive online crowdsourced study of subjective and objective picture quality," *IEEE Transactions on Image Processing*, vol. 25, no. 1, pp. 372–387, 2016.
- [31] Antkowiak, J., Baina, T.J., 2000. Final report from the video quality experts group on the validation of objective models of video quality assessment march. ITU-T Standards Contribution COM.

**AUTHORS**

**Qinglin He** received the B.E. degree in communication system from the School of Communication and Information Engineering, Shanghai University, Shanghai, China, in 2019. She is currently pursuing the M.E. degree in Information and Communication Engineering from Shanghai University. Her research interests include computer vision and image quality assessment.



**Chao Yang** received the B.E. and Ph.D. degree from the School of Communication and Information Engineering, Shanghai University, Shanghai, China, in 2012 and 2017, respectively. From Nov. 2017 to Oct. 2018, he was a Post-Doctoral Fellow with the Department of Electrical and Computer Engineering, University of Southern California (USC), Los Angeles, CA, USA. He joined the School of Communication and Information Engineering, Shanghai University, Shanghai, in 2019, where he is currently a Lecturer. His current research interests include video processing, video compression, and image quality assessment.



**Ping An** received the B.E. and M.E. degrees from the Hefei University of Technology, Hefei, China, in 1990 and 1993, respectively, and the Ph.D. degree from Shanghai University, Shanghai, China, in 2002. In 1993, she joined Shanghai University, where she is currently a Professor with the Video Processing Group, School of Communication and Information Engineering. From 2011 to 2012, she joined the Communication Systems Group, Technische University at Berlin, Germany, as a Visiting Professor. She has finished over 15 projects supported by the National Natural Science Foundation of China, the National Science and Technology Ministry, and the Science and Technology Commission of Shanghai Municipality. Her research interests include image and video processing, with a focus on immersive video processing. She was a recipient of the Second Prize of the Shanghai Municipal Science and Technology Progress Award in 2011, the Second Prize in Natural Sciences of the Ministry of Education in 2016, and the Second Prize in Natural Sciences of the Chinese Institute of Electronics in 2018.

