

# Blind Image Quality Estimation via Distortion Aggravation

Xiongkuo Min<sup>ID</sup>, *Student Member, IEEE*, Guangtao Zhai, *Member, IEEE*, Ke Gu<sup>ID</sup>, *Member, IEEE*, Yutao Liu, *Student Member, IEEE*, and Xiaokang Yang, *Senior Member, IEEE*

**Abstract**—Traditional blind image quality assessment (IQA) measures generally predict quality from a sole distorted image directly. In this paper, we first introduce multiple pseudo reference images (MPRIs) by further degrading the distorted image in several ways and to certain degrees, and then compare the similarities between the distorted image and the MPRIs. Via such distortion aggravation, we can have some references to compare with, i.e., the MPRIs, and utilize the full-reference IQA framework to compute the quality. Specifically, we apply four types and five levels of distortion aggravation to deal with the commonly encountered distortions. Local binary pattern features are extracted to describe the similarities between the distorted image and the MPRIs. The similarity scores are then utilized to estimate the overall quality. More similar to a specific pseudo reference image (PRI) indicates closer quality to this PRI. Owing to the availability of the created multiple PRIs, we can reduce the influence of image content, and infer the image quality more accurately and consistently. Validation is conducted on four mainstream natural scene image and screen content image quality assessment databases, and the proposed method is comparable to or outperforms the state-of-the-art blind IQA measures. The MATLAB source code of the proposed measure will be publicly available.

**Index Terms**—Blind image quality assessment, distortion aggravation, pseudo reference image.

## I. INTRODUCTION

THE QUICK development of network and transmission technologies have boosted various multimedia applications and broadcasting services [1], [2]. With so many easily accessible multimedia services, end-users are expecting better and better quality of experience (QoE) from the service provider. Then how to measure and improve the end-user's

Manuscript received December 1, 2017; revised March 1, 2018; accepted March 6, 2018. Date of publication March 27, 2018; date of current version June 5, 2018. This work was supported by the National Natural Science Foundation of China under Grant 61422112, Grant 61371146, Grant 61521062, and Grant 61527804. (*Corresponding author: Guangtao Zhai.*)

X. Min, G. Zhai, and X. Yang are with the Institute of Image Communication and Network Engineering, Shanghai Key Laboratory of Digital Media Processing and Transmissions, Shanghai Jiao Tong University, Shanghai 200240, China (e-mail: minxiongkuo@gmail.com; zhaiguangtao@sjtu.edu.cn; xkyang@sjtu.edu.cn).

K. Gu is with the Beijing Key Laboratory of Computational Intelligence and Intelligent System, Faculty of Information Technology, Beijing University of Technology, Beijing 100124, China (e-mail: guke.doctor@gmail.com).

Y. Liu is with the School of Computer Science and Technology, Harbin Institute of Technology, Harbin 150001, China (e-mail: liuyutao2008@gmail.com).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TBC.2018.2816783

perceived QoE becomes a urgent problem for both device manufacturer and service provider, and many methods have been proposed to measure and enhance the perceived QoE [3]–[7]. Measuring QoE has aroused increasing attention in recent years, such as QoE evaluation of images and videos [8]–[13]. Among them, blind image quality assessment (IQA) has attracted much interest since it requires no prior knowledge of the original image content and the distortion process, and it can be easily deployed in practical visual communication systems.

Blind or no-reference (NR) IQA is originated from full-reference (FR) IQA, in which the perfect quality reference image is the foundation and the fidelity between it and the distorted image is calculated to measure the quality [14]. FR IQA is reduced to NR IQA to tackle the situation that the reference image may not exist or easy to access. Typically, natural scene statistics (NSS) is modeled, and the distorted image's deviation from the NSS is measured as the quality. Generally, NR IQA measures can be less effective, stable and consistent, since they have to estimate the quality from the sole distorted image, whose characteristics are highly sensitive to the image content. Thus FR IQA measures are preferred if the reference is available.

To have an intuitive understanding of the above phenomenon, we select six NR measures, including DIIVINE [15], BLIINDS2 [16], BRISQUE [17], NFERM [18], IL-NIQE [19], BPRI [20], and six FR measures, including SSIM [21], MS-SSIM [22], VIF [23], GSI [24], FSIM [25], GMSD [26], PSIM [14]. We test their performance on 10% randomly selected commonly distorted (JPEG compression, JPEG2000 compression, Gaussian blur, and white Gaussian noise distortions) images from the TID2013 database [27]. This process is executed for 1,000 times, and the mean performance and standard deviation in terms of Spearman rank-order correlation coefficient (SRCC) are illustrated in Fig. 1. It is observed that FR measures generally have higher performance and lower performance uncertainty (smaller standard deviation) than NR measures, which agrees with the above analyses. Note that the above NR measures are well-trained. If we retrain them using the rest images and test them using the selected images, the uncertainty can be even higher.

Considering the benefits of using the reference, we introduce multiple pseudo reference images (MPRIs) for NR IQA via distortion aggravation. The MPRIs are degraded from the distorted image in several ways and to certain degrees. The

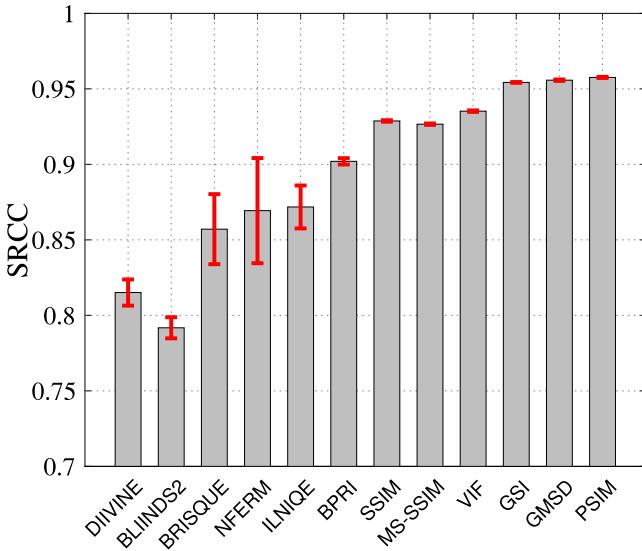


Fig. 1. A comparison of the performance and performance uncertainty of quality measures with and without a reference. The error bar indicates one standard deviation.

pseudo reference image (PRI) is different from the traditional reference image in that it is generated from the distorted image and is assumed to suffer from severer but certain amount of distortion, while the traditional reference image is assumed to have a perfect quality. A systematic introduction of PRI has been given in [20]. We extend PRI to MPRI in this paper. After distortion aggravation and MPRI generation, we follow the FR IQA framework and compare the similarities between the distorted image and the MPRI. More similar to a specific PRI indicates closer quality to this PRI.

We introduce several types of distortion aggravation during MPRI generation to measure the possible distortions existing in the target image. Specifically, we further degrade the distorted image using four types of commonly encountered distortions, including JPEG compression (JPEG), JPEG2000 compression (JP2K), Gaussian blur (GB), and white Gaussian noise (WN), to measure the blocking, ringing, blurring, and noising artifacts. For each type of distortion aggravation, five different but certain levels of distortion is added. A total of twenty PRIs are generated to give references of the same image content degraded by different distortions. Similar to FR IQA, features are then extracted from the distorted image and the PRIs, and the similarities between them are measured as distortion-specific (the same distortion which is used to generate the corresponding PRI) qualities, which are finally fused to an overall general-purpose quality. We name the proposed method as blind MPRI-based (BMPRI) measure.

The rest of this paper is organized as follows. In Section II, we shortly review some related work, including representative blind IQA measures, and several works which follow a similar technique routine as the proposed method. The details of the proposed BMPRI measure are described in Section III. Validation is given in Section IV, which shows that BMPRI is comparable to or outperforms the state-of-the-art blind IQA measures. Section V concludes this paper.

## II. RELATED WORK

### A. Blind IQA Measures

One typical category of blind IQA measures rely on NSS [15]–[17], [19], [28]. These measures follow the same technique routine. First the quality-aware features are extracted in various domains, e.g., wavelet domain [15], DCT domain [16], and spatial domain [17], [19], [28]. Then the statistics of the features extracted from both perfect quality natural images and the distorted image are modeled, respectively. Finally the quality is computed as the distorted image's deviation from natural scenes, and the deviation is described via the modeled statistics. Another typical category of blind IQA measures utilize machine learning techniques, e.g., feature learning [29], [30] and rank learning [31]. Depending on the specific learning techniques, these measures predict quality from the learned or manually extracted quality-aware features. The last typical category of blind IQA measures are motivated by some characteristics of the HVS, e.g., the free energy principle [18], [32], human visual perception of image structures [33], [34]. Some features are designed to simulate these characteristics and these features are then fused to the final quality.

### B. Blind IQA via Distortion Aggravation

All the above blind IQA measures predict quality from single distorted images. A few measures have tried to introduce a “reference” for the distorted image in a NR scenario. Crete *et al.* [35] and Li *et al.* [36] proposed blur quality measures by comparing a blurred image and the re-blurred image. The re-blurring is a distortion aggravation process, and the re-blurred image acts a similar role as the PRI described in this paper, but these two measures can only work for blur estimation. Min *et al.* [20], [37] discussed blind IQA via distortion aggravation systematically, and proposed several PRI-based distortion-specific measures, which are then integrated into a general-purpose measure. In this paper, we improve the method described in [20] from two aspects. First, we generate a series of PRIs with different qualities, rather than one PRI with the worst quality. Second, we unify the feature extraction process and describe all distortions via LBP features, while different features are extracted for different distortions in [20]. With multiple PRIs and a unified feature extraction process, the proposed BMPRI measure can estimate image quality more accurately and consistently. More details of BMPRI are given in the next section.

## III. THE PROPOSED METHOD

As described in Section I, we further degrade the distorted image to generate MPRI, and then measure the similarity between them to predict the quality. Fig. 2 illustrates a framework of the proposed BMPRI measure. Distortion aggravation is introduced in this framework, and we first need to determine the distortion types for this distortion aggravation. Since different distortion types introduce different artifacts, we need to define distortion-specific PRI to be consistent with the characteristics of a given distortion. For example, we can inject

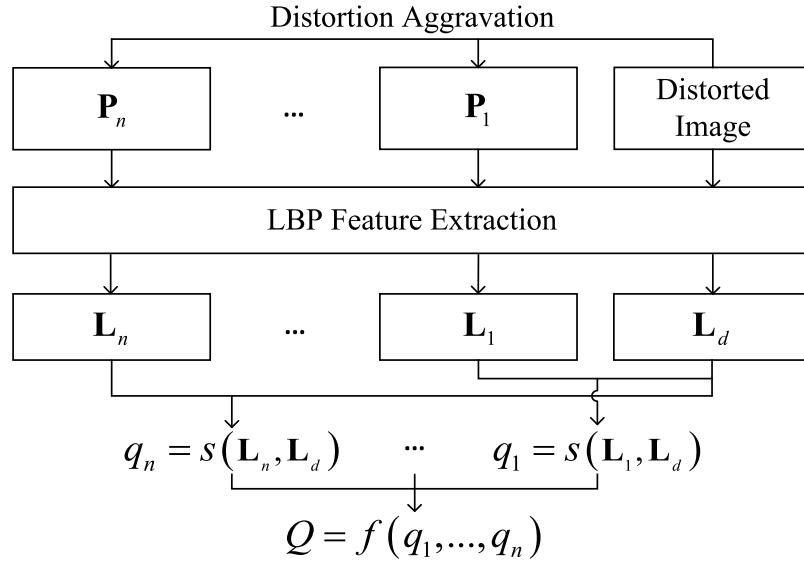


Fig. 2. A framework of the proposed BMPRI measure.  $\mathbf{P}_1$  to  $\mathbf{P}_n$  are the introduced multiple PRIs,  $\mathbf{L}_d$  and  $\mathbf{L}_1$  to  $\mathbf{L}_n$  are the extracted LBP feature maps,  $q_1$  to  $q_n$  indicate the distorted image's similarities to the PRIs, the final quality  $Q$  is fused from  $q_1$  to  $q_n$ .

noise to the distorted image to estimate the noising artifacts, and similarly we can blur the distorted image to estimate the blurring artifacts. Considering that JPEG, JP2K, GB, and WN are the four most commonly encountered distortion types, we further degrade the distorted image using these four distortion types to measure the blocking, ringing, blurring, and noising artifacts. For different types of distortion aggravation, local binary pattern (LBP) features are extracted, and the similarities between the distorted image and MPRIs are used to predict the final quality. The details are as follows.

#### A. Distortion Aggravation

Four types of distortion aggravation are introduced in BMPRI, and for each type, five levels of aggravation are used. To measure the blocking effect, we compress the distorted image  $\mathbf{D}$  to MPRIs  $\mathbf{P}_{k_i}$  using the JPEG encoder

$$\mathbf{P}_{k_i} = \text{JPEG}(\mathbf{D}, Q_i), \quad (1)$$

where  $i$  indicates the  $i$ th level of distortion aggravation, JPEG denotes the JPEG encoder, and  $Q_i$  controls the compression quality. For the ringing effect, we compress the distorted image  $\mathbf{D}$  to MPRIs  $\mathbf{P}_{r_i}$  using the JP2K encoder

$$\mathbf{P}_{r_i} = \text{JP2K}(\mathbf{D}, R_i), \quad (2)$$

JP2K denotes the JP2K encoder, and  $R_i$  controls the compression ratio. For blurring effect, we blur the distorted image  $\mathbf{D}$  to MPRIs  $\mathbf{P}_{b_i}$  using several Gaussian kernels

$$\mathbf{P}_{b_i} = \mathbf{g}_i * \mathbf{D}, \quad (3)$$

where  $*$  is a convolution operator, and  $\mathbf{g}_i$  is a Gaussian kernel with certain standard deviation. For noising effect, we inject noise to the distorted image  $\mathbf{D}$  to get MPRIs  $\mathbf{P}_{n_i}$

$$\mathbf{P}_{n_i} = \mathbf{D} + \mathcal{N}(0, v_i), \quad (4)$$

where  $\mathcal{N}(0, v)$  generates normally distributed random values with 0 mean and  $v$  variance. In Eq. (1)-Eq. (4), the subscripts

$k, r, b, n$  indicate blocking, ringing, blurring, and noising effects, respectively, and five levels of distortion aggravation are introduced for each type, i.e.,  $i = 1, \dots, 5$ .

#### B. LBP Feature Extraction

After distortion aggravation, we compare the similarities between the distorted image and the MPRIs. Note that the PRI is different from the traditional perfect quality reference image. PRI generally has worse quality than the distorted image and it describes the image content under poor quality conditions. Thus traditional FR IQA measures may not be effective enough to describe the similarities between the distorted image and the MPRIs. In this paper, we extract LBP features from the distorted image and the MPRIs, and then measure the feature similarity to predict the quality. LBP is a simple yet powerful visual descriptor widely used in various visual signal processing applications [38]. Quality degradations cause local image structure changes, which can be also captured by LBP.

We compare the center pixel  $g_c$  and its circularly symmetric neighborhoods  $g_p$ , and then binarize their luminance differences and code the binarization results via a numerical value

$$\text{LBP}_{P,R} = \sum_{p=0}^{P-1} u(g_p - g_c), \quad (5)$$

where  $P, R$  denote the neighbour number and radius of the LBP structure,  $u(*)$  is an unit step function

$$u(x) = \begin{cases} 1 & x \geq 0 \\ 0 & x < 0 \end{cases}. \quad (6)$$

The LBP used in our method is different from the non-uniform and uniform definitions given in [38]. We do not differentiate the neighborhoods by attaching a factor  $2^p$ , and do not code the LBP with many spatial transitions as a separate number for

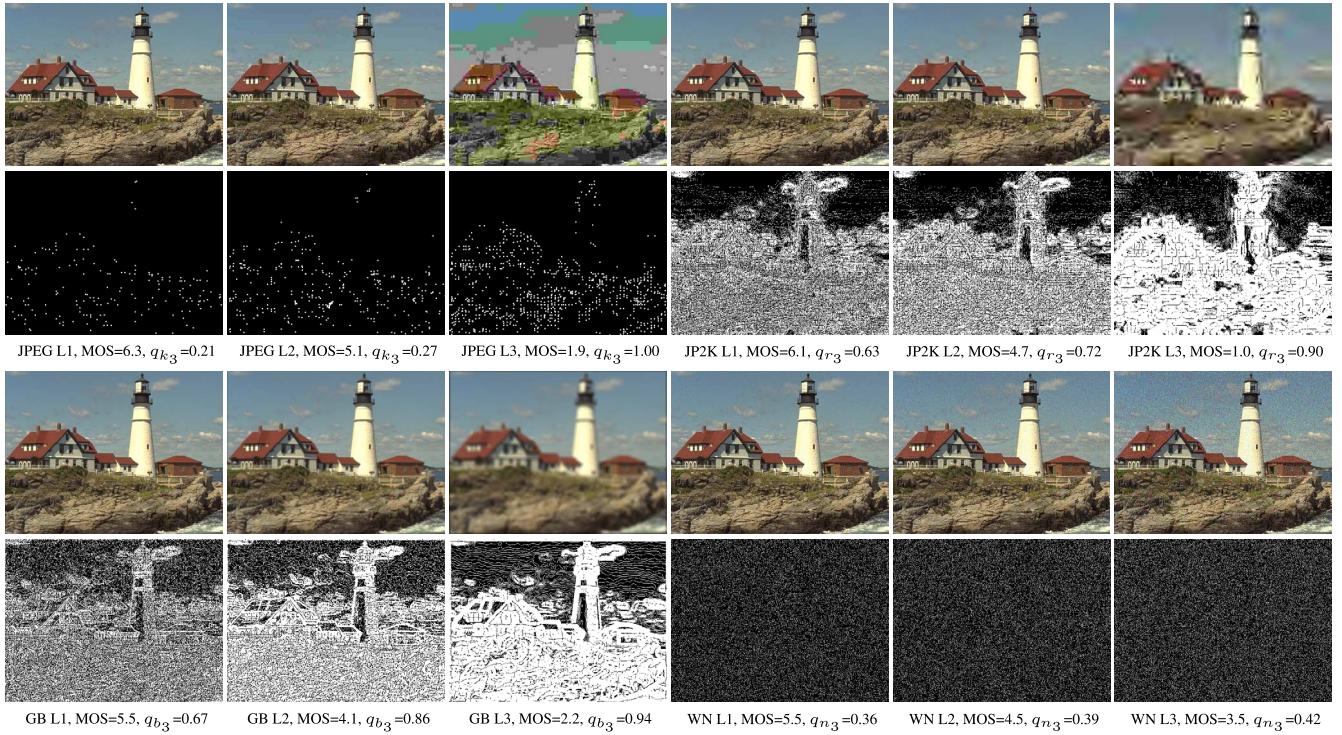


Fig. 3. The overlap between the distorted image's and the MPRI's LBP feature maps for images with different distortion types (JPEG, JP2K, GB, WN) and levels (three levels L1-L3). 1st and 3rd rows: distorted images; 2nd and 4th rows: the corresponding  $\mathbf{L}_o$  maps; MOS: mean opinion score;  $q_{k_3}$ ,  $q_{r_3}$ ,  $q_{b_3}$ ,  $q_{n_3}$ : the similarity scores. More overlap is observed in more distorted images.

simplicity. We set  $P = 4$  and  $R = 1$  for simplicity since it is the simplest one and does not involve any pixel interpolation.

We calculate LBP for all pixels and define the LBP feature map as  $\mathbf{L} = (l_{ij})_{h \times w}$ , whose elements are

$$l_{ij} = \begin{cases} 1 & \text{if } \text{LBP}_{4,1} = c \\ 0 & \text{otherwise} \end{cases}, \quad (7)$$

where  $i, j$  are pixel indexes,  $h \times w$  denotes the image resolution, and  $c = 0, 1, \dots, 4$  indicate five different LBPs. We set  $c$  to different values for different distortion types, since specific LBPs are sensitive to specific distortions. For the distorted image and the MPRIIs, we follow the same feature extraction process, and denote the feature maps as  $\mathbf{L}_d$  and  $\mathbf{L}_m$ , respectively.

### C. Similarities Between the Distorted Image and the MPRIIs

We then compare the similarity between  $\mathbf{L}_d = (l_{dij})_{h \times w}$  and  $\mathbf{L}_m = (l_{mij})_{h \times w}$  to predict the quality. Specifically, we define the overlap between  $\mathbf{L}_d$  and  $\mathbf{L}_m$  as

$$\mathbf{L}_o = (l_{oij})_{h \times w} = (l_{dij} \cdot l_{mij})_{h \times w}. \quad (8)$$

Fig. 3 has illustrated some examples of  $\mathbf{L}_o$  map. Then the similarity between  $\mathbf{L}_d$  and  $\mathbf{L}_m$  can be defined as

$$q = s(\mathbf{L}_d, \mathbf{L}_m) = \frac{\sum_{i,j} l_{oij}}{\sum_{i,j} l_{mij} + 1}, \quad (9)$$

where the numerator and denominator denote the number of non-zero elements in  $\mathbf{L}_o$  and  $\mathbf{L}_m$  maps, respectively. High  $q$  score generally indicates worse quality since the MPRIIs

describe the image content under poor quality conditions. We utilize an average pooling strategy here, but BMPRI may be improved by incorporating visual attention [39]–[41].

### D. Quality Prediction

When measuring the blocking effect, we set  $c$  to 0 and use Eq. (5)–Eq. (9) to compute the similarities between the distorted image  $\mathbf{D}$  and the MPRIIs  $\mathbf{P}_{k_i}$  as  $q_{k_i}$ . We set  $c$  to 0 because this pattern is the most sensitive to the blocking distortion, and  $q$  changes the most significantly with the varying blocking effect of the distorted image. Similarly, we set  $c$  to 2 or 3 to estimate the ringing and blurring effects as  $q_{r_i}$  and  $q_{b_i}$ , and set  $c$  to 0 or 1 to estimate the noising effect as  $q_{n_i}$ . In Fig. 3, we have illustrated some examples of LBP feature similarity maps, the corresponding similarity scores and the mean opinion scores (MOSs). It is observed that the LBP similarity scores have good describing ability for the corresponding distortions, and there is more overlap in more distorted images. We catenate all similarity scores into a twenty dimensional feature vector

$$\mathbf{q} = [q_{k_1}, \dots, q_{k_5}, q_{r_1}, \dots, q_{r_5}, q_{b_1}, \dots, q_{b_5}, q_{n_1}, \dots, q_{n_5}]. \quad (10)$$

Feature vector  $\mathbf{q}$  contains features describing the distorted image's blocking, ringing, blurring, and noising effects, which are the four most frequently encountered artifacts and many practical distortions can be combinations of them. Since distortions in practical applications are often uncertain and the portions of these four distortion effects are also unknown, we

integrate the features in vector  $\mathbf{q}$  into the final quality score  $Q$  via training. Specifically, support vector regression (SVR) is utilized considering its simplicity and high efficiency in regression problems. We use the quality features  $\mathbf{f}_i$  and the corresponding quality labels  $Q_i$  (MOS) of the images in the training set  $\Phi$  to train the *regressor*

$$\text{regressor} = \text{SVR\_TRAIN}(\mathbf{f}_i, Q_i), \quad i \in \Phi, \quad (11)$$

where  $i$  is the image index. After training, we can use the *regressor* to predict the quality of any test image with quality feature  $\mathbf{f}$

$$Q = \text{SVR\_PREDICT}(\mathbf{f}, \text{regressor}). \quad (12)$$

We use the LIBSVM [42] implementation of the SVR with a radial basis function (RBF) kernel. We follow the common SVR parameter settings used in the training of mainstream IQA measures. After training, it can predict the quality of any single images.

#### E. Implementation Details

When applying distortion aggravation, we need to further degrade the distorted image using JPEG, JP2K, GB, and WN distortions, and five degradation levels are used for each type. In Eq. (1), we use the MATLAB implementation of JPEG encoder, and the five quality parameters are from 0 to 8 with a step of 2. In Eq. (2), the five compression ratios are from 150 to 250 with a step of 25. In Eq. (3), five Gaussian kernels with standard deviations from 0.5 to 2.5 with a step of 0.5 are used. In Eq. (4), the five variances are from 0.3 to 0.7 with a step of 0.1. Note that the twenty PRIs have described the image content with poor quality conditions, and they generally have worse quality than the distorted image. We find that the specific quality settings of these PRIs do not influence the overall performance significantly as long as they are in a relative low quality range.

## IV. EXPERIMENTAL RESULTS

Most current IQA measures are designed for natural scene images (NSIs). Since the computer generated screen content has become more and more widespread, screen content image (SCI) quality assessment (QA) is introduced and a few quality measures have been proposed for SCIs [43]–[45]. Traditional NSI quality measures are not effective enough for SCIs, since these measures are highly dependent on NSS and can not describe SCIs well. While specifically designed SCI quality measures have to consider the characteristics of SCIs, and they are not suitable for NSIs. However, in practical visual communication systems, we may encounter both NSIs and SCIs, and we often do not have any prior knowledge about the image types. Efficient general quality measures which are effective for both NSIs and SCIs are highly needed in such applications. The proposed method can fulfill this need and we will compare it with the state-of-the-art blind IQA measures on mainstream NSI and SCI QA databases.

TABLE I  
TEST DATABASES

Type	Name	No. of Ref.	No. of Dist.	Score Type
NSI	LIVE [46]	29	779	DMOS
	TID2013 [27]	25	480	MOS
	CSIQ [47]	30	600	DMOS
SCI	SIQAD [48]	20	560	DMOS

#### A. Experimental Protocol

Four large IQA databases are used as testbeds, including three mainstream NSI QA databases, i.e., LIVE [46], TID2013 [27], CSIQ [47], and one SCI QA database, i.e., SIQAD [48]. The whole LIVE database is used for test, while for the TID2013, CSIQ, and SIQAD databases, we mainly consider the four distortion types which are in common with the LIVE database, i.e., JPEG, JP2K, GB, and WN. An overview of the basic information of the test databases is given in Table I. Besides these common distortions, we will also test the quality measures' generalizability to other non-common distortions.

Ten state-of-the-art blind IQA measures act as competitors in this paper, including: (1) NSS based measures, i.e., DIIIVINE [15], BLIINDS2 [16], BRISQUE [17], NIQE [28], and ILNIQE [19]; (2) learning based measures, i.e., CORNIA [29] and HOSA [30]; (3) human vision based and some other measures, i.e., NFERM [18], LPSI [33], and BPRI [20]. The competitors include both classical and recent measures, and they follow various different technical routines. We think these measures can represent the state-of-the-art in this area.

Following the common practices of IQA model evaluation, we first nonlinearly map the predicted scores using a five-parameter logistic function

$$Q' = \beta_1 \left( \frac{1}{2} - \frac{1}{1 + e^{\beta_2(Q - \beta_3)}} \right) + \beta_4 Q + \beta_5, \quad (13)$$

where  $Q, Q'$  are the predicted and mapped quality scores, respectively;  $\{\beta_i | i = 1, 2, \dots, 5\}$  are parameters determined via curve fitting. Then the consistency between the predicted and ground-truth quality scores is measured to evaluate the IQA model. Specifically, we use the following three consistency metrics: Spearman rank-order correlation coefficient (SRCC), Pearson linear correlation coefficient (PLCC), and root-mean-square error (RMSE), which measure the prediction monotonicity, linearity, and accuracy, respectively.

#### B. Overall Performance Comparison

Following the common practices of training based IQA model evaluation [15]–[18], we split the image database into two completely separate sets: a training set with 80% distorted images and a testing set with the rest 20% distorted images. The distorted images corresponding to the same reference image are assigned to the same set to ensure a complete separation of the training and testing image content. For DIIIVINE, BLIINDS-II, BRISQUE, NFERM, and the proposed method,

TABLE II  
OVERALL MEDIAN SRCC, PLCC, AND RMSE PERFORMANCE COMPARISON

Database	Criteria	DIIVINE	BLLIINDS2	BRISQUE	NIQE	ILNIQE	CORNIA	HOSA	NFERM	LPSI	BPRI	BMPRI
LIVE	SRCC	0.8729	0.9109	<b>0.9390</b>	0.9102	0.9046	-	-	0.9352	0.8199	0.9047	0.9310
	PLCC	0.8828	0.9250	<b>0.9439</b>	0.9088	0.9068	-	-	0.9401	0.8323	0.9059	0.9329
	RMSE	12.837	10.297	<b>9.0156</b>	11.382	11.512	-	-	9.2892	15.096	6.2055	9.8335
TID2013	SRCC	0.7498	0.8580	0.8542	0.8111	0.8777	0.8938	0.9021	0.9078	0.7156	0.8991	<b>0.9287</b>
	PLCC	0.7989	0.8957	0.8883	0.8243	0.8930	0.9067	0.9223	0.9308	0.8258	0.8917	<b>0.9466</b>
	RMSE	0.8379	0.6215	0.6430	0.7888	0.6282	0.5901	0.5409	0.5100	0.7901	0.6320	<b>0.4488</b>
CSIQ	SRCC	0.8573	0.8917	0.8699	0.8837	0.8867	-	-	<b>0.9140</b>	0.7772	0.9034	0.9085
	PLCC	0.8946	0.9221	0.8991	0.9050	0.9190	-	-	<b>0.9434</b>	0.8742	0.9251	0.9339
	RMSE	0.1258	0.1087	0.1240	0.1177	0.1082	-	-	<b>0.0932</b>	0.1354	0.1071	0.0989
SIQAD	SRCC	0.7203	0.7278	0.7071	0.4982	0.5446	0.5267	0.3852	0.7647	0.4288	<b>0.7825</b>	0.7717
	PLCC	0.7536	0.7766	0.7588	0.5433	0.6041	0.5560	0.4556	0.8049	0.5406	0.8092	<b>0.8129</b>
	RMSE	9.1101	8.8384	9.0712	11.583	11.055	11.447	12.328	8.2414	11.672	8.1668	<b>8.0417</b>
Average	SRCC	0.8001	0.8491	0.8425	0.7758	0.8034	0.7976	0.7620	0.8804	0.6854	0.8724	<b>0.8850</b>
	PLCC	0.8325	0.8798	0.8725	0.7953	0.8307	0.8118	0.7924	0.9048	0.7682	0.8830	<b>0.9066</b>
#Hit		0	0	3	0	0	0	0	3	0	1	7

TABLE III  
MEDIAN SRCC PERFORMANCE COMPARISON ON INDIVIDUAL DISTORTIONS

Database	Criteria	DIIVINE	BLLIINDS2	BRISQUE	NIQE	ILNIQE	CORNIA	HOSA	NFERM	LPSI	BPRI	BMPRI
LIVE	JPEG	0.8854	0.9458	0.9646	0.9436	0.9461	-	-	0.9647	0.9668	<b>0.9677</b>	0.9668
	JP2K	0.8193	0.9325	0.9128	0.9265	0.9036	-	-	0.9371	0.9372	0.9195	<b>0.9393</b>
	GB	0.8714	0.9083	<b>0.9504</b>	0.9408	0.9293	-	-	0.9132	0.9266	0.9348	0.9181
	WN	0.9591	0.9444	0.9791	0.9711	0.9795	-	-	0.9844	0.9582	0.9835	<b>0.9860</b>
	FF	0.8053	0.8527	<b>0.8781</b>	0.8643	0.8458	-	-	0.8534	0.7820	0.8196	0.8269
TID2013	JPEG	0.6892	0.8149	0.8312	0.8723	0.8738	0.9162	0.9179	0.8877	0.9246	0.9177	<b>0.9262</b>
	JP2K	0.7860	0.9062	0.8652	0.8991	0.9108	0.9031	<b>0.9300</b>	0.9106	0.9054	0.8853	0.9262
	GB	0.8585	0.8922	0.8705	0.8199	0.8466	<b>0.9346</b>	0.9177	0.8923	0.8835	0.8712	0.9162
	WN	0.6633	0.6762	0.8254	0.8530	0.8923	0.7815	0.8485	0.8946	0.8208	0.9334	<b>0.9338</b>
CSIQ	JPEG	0.8825	0.9035	0.9018	0.8830	0.9039	-	-	0.9158	<b>0.9539</b>	0.9332	0.9181
	JP2K	0.8545	0.8903	0.8381	0.9240	0.9221	-	-	0.9079	<b>0.9226</b>	0.8750	0.8999
	GB	0.8751	0.9088	0.8913	0.9032	0.8688	-	-	<b>0.9292</b>	0.9168	0.9075	0.9180
	WN	0.8009	0.8657	0.9071	0.8325	0.8695	-	-	0.9074	0.7244	<b>0.9399</b>	0.9275
SIQAD	JPEG	0.4792	0.4028	0.5419	0.5287	0.3651	0.1823	0.4691	0.6163	<b>0.7737</b>	0.7447	0.6180
	JP2K	0.5331	0.6527	0.3591	0.2909	0.4540	0.6500	0.5260	0.5041	0.5914	0.6593	<b>0.6853</b>
	GB	0.8013	0.8177	0.8202	0.6212	0.5769	0.7334	0.2365	<b>0.8922</b>	0.7808	0.8539	0.8243
	WN	0.7967	0.8615	0.8196	0.8473	0.8539	0.7567	0.7455	0.8333	<b>0.8815</b>	0.8730	0.8591
Average		0.7859	0.8339	0.8327	0.8189	0.8201	0.7322	0.6989	0.8673	0.8618	<b>0.8835</b>	0.8817
#Hit		0	0	2	0	0	1	1	2	4	3	<b>5</b>

we retrain them on the training set and test them on the testing set. While for the rest methods, we test them on the same testing set for fair comparison. We repeat this training-testing process for 1,000 times, and the median SRCC, PLCC, and RMSE performance is listed in Table II.

The databases listed in Table I are used as testbeds, and the mean performance on these four databases is also reported. Since CORNIA and HOSA are trained on the LIVE database and the images from the CSIQ database are used to construct the codebook, their performance on these two databases is not reported in Table II. It is observed that the proposed BMPRI method is comparable to the state-of-the-art blind IQA measures, and it has the best performance from an averaging perspective. Another observation is that most blind IQA measures undergo performance drop when transferring from NSIs to SCIs. While BMPRI and BPRI are the two methods with

the least performance drop, which suggests that the proposed method has good content type generalizability.

### C. Performance on Single Distortions

Besides the overall performance on individual databases, we also evaluate all blind IQA measures on individual distortions. The same training-testing procedures described in Section IV-B are conducted. The 80% distorted images belonging to the training set are all used to train the models, but for the rest 20% distorted images belonging to the testing set, only images degraded by the target distortion type are used for testing. The median performance is listed in Table III. Only SRCC is reported for simplicity, but similar evaluation results can be obtained using other evaluation criteria. It is observed that the proposed BMPRI measure is also comparable to the

TABLE IV  
SRCC PERFORMANCE ON SINGLE DISTORTIONS OF THE TID2013 DATABASE

Distortions	DIIVINE	BLIINDS2	BRISQUE	NIQE	ILNIQE	CORNIA	HOSA	NFERM	LPSI	BPRI	BMPRI
#01	0.8553	0.6468	0.8520	0.8187	0.8767	0.7354	0.8653	0.8581	0.7690	0.9181	0.8477
#02	0.7120	0.4762	0.7089	0.6701	0.8159	0.7076	0.7687	0.7096	0.4955	0.8587	0.7831
#03	0.4626	0.5862	0.4916	0.6659	0.9233	0.6892	0.5804	0.2184	0.6968	0.5293	0.5516
#04	0.6752	0.6183	0.5767	0.7464	0.5134	0.7141	0.7250	0.2210	0.0462	0.7479	0.8061
#05	0.8778	0.7229	0.7526	0.8454	0.8691	0.7972	0.8642	0.8813	0.9250	0.9263	0.8520
#06	0.8063	0.6525	0.6289	0.7446	0.7556	0.7634	0.7878	0.1728	0.4324	0.4585	0.3965
#07	0.1650	0.7370	0.7932	0.8514	0.8721	0.0922	0.7991	0.7747	0.8537	0.4898	0.7247
#08	0.8344	0.8367	0.8137	0.7986	0.8148	0.9274	0.9046	0.8501	0.8408	0.8593	0.8954
#09	0.7231	0.6884	0.5849	0.5900	0.7494	0.8459	0.0690	0.6389	0.2487	0.4210	0.4359
#10	0.6288	0.8360	0.8448	0.8468	0.8340	0.8958	0.9115	0.8722	0.9123	0.9107	0.9044
#11	0.8534	0.8883	0.8927	0.8890	0.8583	0.9009	0.9194	0.8097	0.8988	0.8680	0.9022
#12	0.2387	0.1098	0.3163	0.0006	0.2819	0.6991	0.7085	0.1322	0.0911	0.7887	0.5524
#13	0.0606	0.6409	0.3595	0.5114	0.5240	0.6762	0.3689	0.1684	0.6106	0.4883	0.4688
#14	0.0598	0.0997	0.1459	0.0682	0.0808	0.2332	0.3714	0.0646	0.0520	0.0086	0.0561
#15	0.0928	0.2440	0.2233	0.1218	0.1334	0.2287	0.2912	0.2020	0.1372	0.2333	0.1759
#16	0.0104	0.0963	0.1241	0.1639	0.1840	0.0844	0.0843	0.0213	0.3409	0.1106	0.2315
#17	0.4601	0.0011	0.0404	0.0171	0.0136	0.1814	0.1447	0.2178	0.1992	0.1846	0.0095
#18	0.0684	0.0119	0.1126	0.2481	0.1655	0.0353	0.0675	0.3067	0.3018	0.3786	0.3766
#19	0.7873	0.6193	0.7242	0.6934	0.6936	0.6574	0.7882	0.7162	0.6959	0.8612	0.7843
#20	0.1156	0.1663	0.0076	0.1544	0.3614	0.5235	0.3589	0.1427	0.0181	0.0691	0.3785
#21	0.6327	0.4552	0.6856	0.8023	0.8287	0.8654	0.8513	0.6541	0.2356	0.5977	0.7435
#22	0.4362	0.7677	0.7652	0.7881	0.7504	0.3919	0.7562	0.4790	0.8998	0.6753	0.7290
#23	0.6608	0.6445	0.6166	0.5671	0.6793	0.8183	0.6998	0.6430	0.6953	0.7253	0.7495
#24	0.8334	0.8257	0.7841	0.8340	0.8643	0.8536	0.7610	0.7847	0.8620	0.7873	0.7654
Average	0.5021	0.5155	0.5352	0.5599	0.6018	0.5966	0.6020	0.4808	0.5108	0.5790	0.5884

state-of-the-art methods when evaluated on single distortions, which agrees with the overall performance evaluation results described in Section IV-B.

#### D. Generalizability to Other Distortions

To test the blind IQA measures' generalizability to other distortions, we evaluate them on all 24 single distortions of the entire TID2013 database, including #01 additive Gaussian noise, #02 additive noise in color components, #03 spatially correlated noise, #04 masked noise, #05 high frequency noise, #06 impulse noise, #07 quantization noise, #08 Gaussian blur, #09 image denoising, #10 JPEG compression, #11 JPEG2000 compression, #12 JPEG transmission errors, #13 JPEG2000 transmission errors, #14 non eccentricity pattern noise, #15 local block-wise distortions, #16 mean shift, #17 contrast change, #18 change of color saturation, #19 multiplicative Gaussian noise, #20 comfort noise, #21 lossy compression of noisy images, #22 color quantization with dither, #23 chromatic aberrations, and #24 sparse sampling and reconstruction.

For all competitors, we use the original implementations released by the authors. Most training based measures are well trained on the LIVE database, and we use them directly. For the proposed method, we also train it on the LIVE database and test it on the TID2013 database. The SRCC performance is summarized in Table IV. The proposed method is comparable to the state-of-the-art measures from an averaging perspective. Note that the generalizability of the proposed method is slightly inferior to the best performing measures like ILNIQE, CORNIA and HOSA. It is probably because we only introduce four types of distortion aggravation, i.e., the four common distortion types, when deriving the MPRIs. For

TABLE V  
COMPUTATIONAL COMPLEXITY

Method	Time (seconds/image)
DIIVINE	8.0033
BLIINDS2	16.0604
BRISQUE	0.4493
NIQE	0.0964
ILNIQE	3.5990
CORNIA	2.9422
HOSA	0.2519
NFERM	19.3261
LPSI	0.0126
BPRI	0.7894
BMPRI	1.3965

other distortions which are distinctive from these distortions, the proposed method may encounter slight generalizability problems. But for such distortions, the state-of-the-art measures can not handle them effectively either, for example the mean shift and contrast change distortions.

#### E. Computational Complexity

To analyse the computational complexity of all blind IQA measures, we test them on 100 images with a fixed resolution of  $512 \times 512$ , and report the average running time (seconds/image). The experiment is conducted on a computer with 4.20 GHz Intel Core i7-7700K CPU and 16 GB RAM. We use the implementations released by the original authors. The running time includes all feature extraction and regression time, and the results are summarized in Table V. Though the proposed method does not have the shortest running time, it still has considerable low computational complexity. Note that

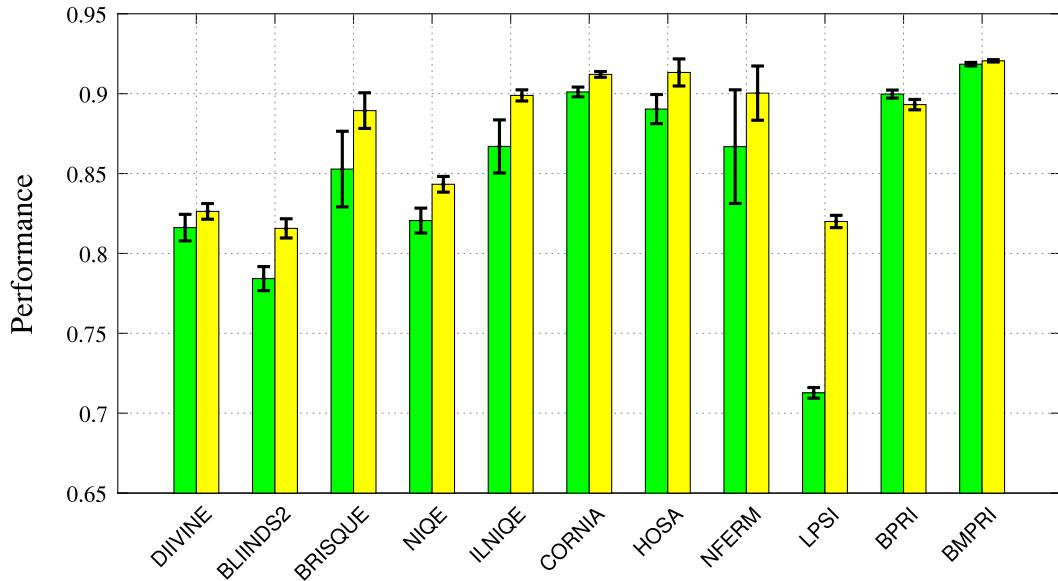


Fig. 4. Mean and standard error bar of the SRCC and PLCC values obtained from the 1,000 random tests on the TID2013 database.

we have not optimized and accelerated the code yet. The most time consuming operations are the JPEG and JPEG2000 compression introduced during distortion aggravation, while there are many methods and solutions to accelerate such operations in practical use.

#### F. Stability and Performance Uncertainty

As described in Section I and illustrated in Fig. 1, most blind IQA measures are not stable since they have to estimate quality from a single distorted image, whose characteristics are highly sensitive to the image content. Similar to the test conducted in Fig. 1, we test all blind IQA measures' stability by evaluating them on the four common distortions, i.e., JPEG, JP2K, GB, and WN, which are included in the TID2013 database. We randomly select 10% images from the overall 480 distorted images, and test their performance. The image selection criteria is similar to the image selection in the training-testing process described in Section IV-B. This random selecting and testing are executed for 1,000 times, and the mean performance and standard deviation in terms of SRCC and PLCC are illustrated in Fig. 4.

We have two observations. First, the proposed BMPRI method shows the best performance, which agrees with the previous validations given in Section IV-B. Second, BMPRI has the smallest standard deviations in terms of both SRCC and PLCC. It indicates that BMPRI is more stable and consistent compared with the state-of-the-art. Most of current blind IQA measures have high performance uncertainty because of their NR nature. These measures rely on NSS, and they predict quality by measuring one single image's deviation from the NSS. While a single image's characteristics can be very sensitive to the image content, thus these measures can be more effective on some image content and less effective on some other image content. As shown in Fig. 1, FR measures suffer less from this problem since they mainly predict quality by comparing two images, and the influence of image content

is reduced during such comparison process. The proposed method inherits this merit since we also predict quality by comparing image pairs using the FR IQA framework.

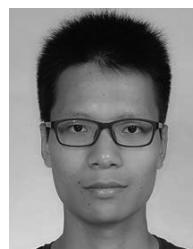
## V. CONCLUSION

In this paper, we introduce the FR IQA framework into NR IQA. Since the reference image is missing in the NR scenario, we create MPRI via distortion aggravation. We further degrade the distorted image to the MPRI in several ways and to certain degrees. Then the MPRI act as references, and we measure the similarities between the distorted image and the MPRI. More similar to a specific PRI indicates closer quality to this PRI. The similarity scores are then utilized to predict the final quality. Specifically, we utilize four types and five levels of distortion aggravation on the distorted image to generate twenty PRIs, and LBP features are extracted to measure the similarities between the distorted image and these twenty PRIs. Experimental results on four large IQA databases have verified the effectiveness and efficiency of the proposed method. On one hand, the proposed method is comparable to the state-of-the-art blind IQA measures in terms of performance. On the other hand, it is more stable and consistent than the state-of-the-art, since we predict quality by comparing image pairs and the influence of image content is significantly reduced via such image comparison.

## REFERENCES

- [1] L. Jalal and M. Murroni, "Enhancing TV broadcasting services: A survey on multimedia quality of experience," in *Proc. IEEE Int. Symp. Broadband Multimedia Syst. Broadcast.*, 2017, pp. 1–7.
- [2] L. Atzori, D. D. Giusto, and M. Murroni, "Performance analysis of fractal modulation transmission over fast-fading wireless channels," *IEEE Trans. Broadcast.*, vol. 48, no. 2, pp. 103–110, Jun. 2002.
- [3] Z. Luo, L. Song, S. Zheng, and N. Ling, "H.264/advanced video control perceptual optimization coding based on JND-directed coefficient suppression," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 23, no. 6, pp. 935–948, Jun. 2013.

- [4] J. Zou *et al.*, "Prioritized flow optimization with multi-path and network coding based routing for scalable multirate multicasting," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 21, no. 3, pp. 259–273, Mar. 2011.
- [5] X. Liu *et al.*, "Sparsity-based image error concealment via adaptive dual dictionary learning and regularization," *IEEE Trans. Image Process.*, vol. 26, no. 2, pp. 782–796, Feb. 2017.
- [6] X. Liu, D. Zhao, J. Zhou, W. Gao, and H. Sun, "Image interpolation via graph-based Bayesian label propagation," *IEEE Trans. Image Process.*, vol. 23, no. 3, pp. 1084–1096, Mar. 2014.
- [7] J. Zhou, O. C. Au, G. Zhai, Y. Y. Tang, and X. Liu, "Scalable compression of stream cipher encrypted images through context-adaptive sampling," *IEEE Trans. Inf. Forensics Security*, vol. 9, no. 11, pp. 1857–1868, Nov. 2014.
- [8] K. Gu *et al.*, "Analysis of distortion distribution for pooling in image quality prediction," *IEEE Trans. Broadcast.*, vol. 62, no. 2, pp. 446–456, Jun. 2016.
- [9] M. Liu, K. Gu, G. Zhai, P. Le Callet, and W. Zhang, "Perceptual reduced-reference visual quality assessment for contrast alteration," *IEEE Trans. Broadcast.*, vol. 63, no. 1, pp. 71–81, Mar. 2017.
- [10] L. A. da Silva Cruz, M. Cordina, C. J. Debono, and P. A. A. Assuncao, "Quality monitor for 3-D video over hybrid broadcast networks," *IEEE Trans. Broadcast.*, vol. 62, no. 4, pp. 785–799, Dec. 2016.
- [11] Q. Wu, H. Li, F. Meng, and K. N. Ngan, "Toward a blind quality metric for temporally distorted streaming video," *IEEE Trans. Broadcast.*, to be published.
- [12] R. Sotelo, J. Joskowicz, M. Anedda, M. Murroni, and D. D. Giusto, "Subjective video quality assessments for 4K UHDTV," in *Proc. IEEE Int. Symp. Broadband Multimedia Syst. Broadcast.*, 2017, pp. 1–6.
- [13] L. Song, X. Tang, W. Zhang, X. Yang, and P. Xia, "The SJTU 4K video sequence dataset," in *Proc. Int. Workshop Qual. Multimedia Exp.*, Klagenfurt, Austria, 2013, pp. 34–35.
- [14] K. Gu, L. Li, H. Lu, X. Min, and W. Lin, "A fast reliable image quality predictor by fusing micro- and macro-structures," *IEEE Trans. Ind. Electron.*, vol. 64, no. 5, pp. 3903–3912, May 2017.
- [15] A. K. Moorthy and A. C. Bovik, "Blind image quality assessment: From natural scene statistics to perceptual quality," *IEEE Trans. Image Process.*, vol. 20, no. 12, pp. 3350–3364, Dec. 2011.
- [16] M. A. Saad, A. C. Bovik, and C. Charrier, "Blind image quality assessment: A natural scene statistics approach in the DCT domain," *IEEE Trans. Image Process.*, vol. 21, no. 8, pp. 3339–3352, Aug. 2012.
- [17] A. Mittal, A. K. Moorthy, and A. C. Bovik, "No-reference image quality assessment in the spatial domain," *IEEE Trans. Image Process.*, vol. 21, no. 12, pp. 4695–4708, Dec. 2012.
- [18] K. Gu, G. Zhai, X. Yang, and W. Zhang, "Using free energy principle for blind image quality assessment," *IEEE Trans. Multimedia*, vol. 17, no. 1, pp. 50–63, Jan. 2015.
- [19] L. Zhang, L. Zhang, and A. C. Bovik, "A feature-enriched completely blind image quality evaluator," *IEEE Trans. Image Process.*, vol. 24, no. 8, pp. 2579–2591, Aug. 2015.
- [20] X. Min *et al.*, "Blind quality assessment based on pseudo reference image," *IEEE Trans. Multimedia*, to be published.
- [21] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.
- [22] Z. Wang, E. P. Simoncelli, and A. C. Bovik, "Multiscale structural similarity for image quality assessment," in *Proc. IEEE Asilomar Conf. Signal., Syst., Comput.*, vol. 2. Pacific Grove, CA, USA, 2003, pp. 1398–1402.
- [23] H. R. Sheikh and A. C. Bovik, "Image information and visual quality," *IEEE Trans. Image Process.*, vol. 15, no. 2, pp. 430–444, Feb. 2006.
- [24] A. Liu, W. Lin, and M. Narwaria, "Image quality assessment based on gradient similarity," *IEEE Trans. Image Process.*, vol. 21, no. 4, pp. 1500–1512, Apr. 2012.
- [25] L. Zhang, L. Zhang, X. Mou, and D. Zhang, "FSIM: A feature similarity index for image quality assessment," *IEEE Trans. Image Process.*, vol. 20, no. 8, pp. 2378–2386, Aug. 2011.
- [26] W. Xue, L. Zhang, X. Mou, and A. C. Bovik, "Gradient magnitude similarity deviation: A highly efficient perceptual image quality index," *IEEE Trans. Image Process.*, vol. 23, no. 2, pp. 684–695, Feb. 2014.
- [27] N. Ponomarenko *et al.*, "Image database TID2013: Peculiarities, results and perspectives," *Signal Process. Image Commun.*, vol. 30, pp. 57–77, Jan. 2015.
- [28] A. Mittal, R. Soundararajan, and A. C. Bovik, "Making a 'completely blind' image quality analyzer," *IEEE Signal Process. Lett.*, vol. 20, no. 3, pp. 209–212, Mar. 2013.
- [29] P. Ye, J. Kumar, L. Kang, and D. Doermann, "Unsupervised feature learning framework for no-reference image quality assessment," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Providence, RI, USA, 2012, pp. 1098–1105.
- [30] J. Xu *et al.*, "Blind image quality assessment based on high order statistics aggregation," *IEEE Trans. Image Process.*, vol. 25, no. 9, pp. 4444–4457, Sep. 2016.
- [31] K. Ma, W. Liu, T. Liu, Z. Wang, and D. Tao, "dipIQ: Blind image quality assessment by learning-to-rank discriminable image pairs," *IEEE Trans. Image Process.*, vol. 26, no. 8, pp. 3951–3964, Aug. 2017.
- [32] G. Zhai, X. Wu, X. Yang, W. Lin, and W. Zhang, "A psychovisual quality metric in free-energy principle," *IEEE Trans. Image Process.*, vol. 21, no. 1, pp. 41–52, Jan. 2012.
- [33] Q. Wu, Z. Wang, and H. Li, "A highly efficient method for blind image quality assessment," in *Proc. IEEE Int. Conf. Image Process.*, Quebec City, QC, Canada, 2015, pp. 339–343.
- [34] Q. Li, W. Lin, J. Xu, and Y. Fang, "Blind image quality assessment using statistical structural and luminance features," *IEEE Trans. Multimedia*, vol. 18, no. 12, pp. 2457–2469, Dec. 2016.
- [35] F. Crete, T. Dolmire, P. Ladret, and M. Nicolas, "The blur effect: Perception and estimation with a new no-reference perceptual blur metric," in *Proc. SPIE*, vol. 6492, Feb. 2007, pp. 1–11.
- [36] C. Li, W. Yuan, A. C. Bovik, and X. Wu, "No-reference blur index using blur comparisons," *IET Electron. Lett.*, vol. 47, no. 17, pp. 962–963, Aug. 2011.
- [37] X. Min *et al.*, "Blind quality assessment of compressed images via pseudo structural similarity," in *Proc. IEEE Int. Conf. Multimedia Expo*, Seattle, WA, USA, 2016, pp. 1–6.
- [38] T. Ojala, M. Pietikainen, and T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 7, pp. 971–987, Jul. 2002.
- [39] X. Min, G. Zhai, Z. Gao, and K. Gu, "Visual attention data for image quality assessment databases," in *Proc. IEEE Int. Symp. Circuits Syst.*, Melbourne, VIC, Australia, 2014, pp. 894–897.
- [40] X. Min, G. Zhai, K. Gu, and X. Yang, "Fixation prediction through multimodal analysis," *ACM Trans. Multimedia Comput. Commun. Appl.*, vol. 13, no. 1, pp. 1–23, 2017.
- [41] X. Min *et al.*, "Visual attention analysis and prediction on human faces," *Inf. Sci.*, vol. 420, pp. 417–430, Dec. 2017.
- [42] C.-C. Chang and C.-J. Lin, "LIBSVM: A library for support vector machines," *ACM Trans. Intell. Syst. Technol.*, vol. 2, no. 3, p. 27, 2011.
- [43] X. Min *et al.*, "Unified blind quality assessment of compressed natural, graphic, and screen content images," *IEEE Trans. Image Process.*, vol. 26, no. 11, pp. 5462–5474, Nov. 2017.
- [44] X. Min, K. Gu, G. Zhai, M. Hu, and X. Yang, "Saliency-induced reduced-reference quality index for natural scene and screen content images," *Signal Process.*, vol. 145, pp. 127–136, Apr. 2018.
- [45] K. Gu *et al.*, "Evaluating quality of screen content images via structural variation analysis," *IEEE Trans. Vis. Comput. Graphics*, to be published.
- [46] H. R. Sheikh, Z. Wang, L. Cormack, and A. C. Bovik, *LIVE Image Quality Assessment Database Release 2*. Accessed: Mar. 22, 2014. [Online]. Available: <http://live.ece.utexas.edu/research/quality>
- [47] E. C. Larson and D. M. Chandler, "Most apparent distortion: Full-reference image quality assessment and the role of strategy," *J. Electron. Imag.*, vol. 19, no. 1, 2010, Art. no. 011006.
- [48] H. Yang, Y. Fang, and W. Lin, "Perceptual quality assessment of screen content images," *IEEE Trans. Image Process.*, vol. 24, no. 11, pp. 4408–4421, Nov. 2015.



**Xiongkuo Min** received the B.E. degree from Wuhan University, Wuhan, China, in 2013. He is currently pursuing the Ph.D. degree with the Institute of Image Communication and Network Engineering, Shanghai Jiao Tong University, Shanghai, China. From 2016 to 2017, he was a Visiting Student with the Department of Electrical and Computer Engineering, University of Waterloo, Canada. His research interests include image quality assessment, visual attention modeling, and perceptual signal processing. He was a recipient of the Best Student Paper Award of ICME 2016.



**Guangtao Zhai** (M'10) received the B.E. and M.E. degrees from Shandong University, Shandong, China, in 2001 and 2004, respectively, and the Ph.D. degree from Shanghai Jiao Tong University, Shanghai, China, in 2009, where he is currently a Research Professor with the Institute of Image Communication and Information Processing. From 2008 to 2009, he was a Visiting Student with the Department of Electrical and Computer Engineering, McMaster University, Hamilton, ON, Canada, where he was a Post-Doctoral Fellow from 2010 to 2012.

From 2012 to 2013, he was a Humboldt Research Fellow with the Institute of Multimedia Communication and Signal Processing, Friedrich Alexander University of Erlangen-Nuremberg, Germany. His research interests include multimedia signal processing and perceptual signal processing. He was a recipient of the National Excellent Ph.D. Thesis Award from the Ministry of Education of China in 2012.



**Ke Gu** received the B.S. and Ph.D. degrees in electronic engineering from Shanghai Jiao Tong University, Shanghai, China, in 2009 and 2015, respectively. His research interests include quality assessment, contrast enhancement, visual saliency detection, and air quality prediction. He was a recipient of the Best Paper Award at the IEEE International Conference on Multimedia and Expo in 2016, and the Excellent Ph.D. Thesis Award from the Chinese Institute of Electronics in 2016. He is an Associate Editor for the IEEE ACCESS, and

a Reviewer for the IEEE TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS, the IEEE TRANSACTIONS ON IMAGE PROCESSING, the IEEE TRANSACTIONS ON CYBERNETICS, the IEEE TRANSACTIONS ON INDUSTRIAL ELECTRONICS, the IEEE TRANSACTIONS ON MULTIMEDIA, the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, the IEEE TRANSACTIONS ON BROADCASTING, J-STSP, SPL, ACCESS, *Information Sciences*, *Neurocomputing*, SPIC, JVCI, DSP, MTAP, and ELL. He has reviewed over 50 journal papers each year. He is a leading Special Session Organizer in VCIP 2016 and ICIP 2017.



**Yutao Liu** received the B.S. and M.S. degrees in computer science from the Harbin Institute of Technology, Harbin, China, in 2011 and 2013, respectively, where he is currently pursuing the Ph.D. degree with the School of Computer Science and Technology. From 2014 to 2016, he was a Research Assistant with the Institute of Image Communication and Network Engineering, Shanghai Jiao Tong University, Shanghai, China. His current research interests include image processing and image quality assessment.



**Xiaokang Yang** (M'00–SM'04) received the B.S. degree from Xiamen University, Xiamen, China, in 1994, the M.S. degree from the Chinese Academy of Sciences, Shanghai, China, in 1997, and the Ph.D. degree from Shanghai Jiao Tong University, Shanghai, in 2000.

He is currently a Distinguished Professor with the School of Electronic Information and Electrical Engineering, and the Deputy Director of the Institute of Image Communication and Information Processing, Shanghai Jiao Tong University. From

2000 to 2002, he was a Research Fellow with the Centre for Signal Processing, Nanyang Technological University, Singapore. From 2002 to 2004, he was a Research Scientist with the Institute for Infocomm Research, Singapore. From 2007 to 2008, he was an Alexander von Humboldt Research Fellow with the Institute for Computer Science, University of Freiburg, Freiburg im Breisgau, Germany. He has published over 200 refereed papers, and holds 60 patents. His current research interests include image processing and communication, computer vision, and machine learning.

Prof. Yang is an Associate Editor of the IEEE TRANSACTIONS ON MULTIMEDIA and a Senior Associate Editor of the *IEEE Signal Processing Letters*. He was a Series Editor of Springer CCIS, and an Editorial Board Member of *Digital Signal Processing*. He is also the Chair of the Multimedia Big Data Interest Group of MMTC Technical Committee, IEEE Communication Society. He is a member of the Asia-Pacific Signal and Information Processing Association, the VSPC Technical Committee of the IEEE Circuits and Systems Society, and the MMSP Technical Committee of the IEEE Signal Processing Society.