Blind Quality Assessment Based on Pseudo-Reference Image

Xiongkuo Min, *Student Member, IEEE*, Ke Gu, *Member, IEEE*, Guangtao Zhai, *Member, IEEE*, Jing Liu, Xiaokang Yang, *Senior Member, IEEE*, and Chang Wen Chen, *Fellow, IEEE*

***Abstract*—Traditional full-reference image quality assessment (IQA) metrics generally predict the quality of the distorted image by measuring its deviation from a perfect quality image called reference image. When the reference image is not fully available, the reduced-reference and no-reference IQA metrics may still be able to derive some characteristics of the perfect quality images, and then measure the distorted image's deviation from these characteristics. In this paper, contrary to the conventional IQA metrics, we utilize a new "reference" called pseudo-reference image (PRI) and a PRI-based blind IQA (BIQA) framework. Different from a traditional reference image, which is assumed to have a perfect quality, PRI is generated from the distorted image and is assumed to suffer from the severest distortion for a given application. Based on the PRI-based BIQA framework, we develop distortion-specific metrics to estimate blockiness, sharpness, and noisiness. The PRI-based metrics calculate the similarity between the distorted image's and the PRI's structures. An image suffering from severer distortion has a higher degree of similarity with the corresponding PRI. Through a two-stage quality regression after a distortion identification framework, we then integrate the PRI-based distortion-specific metrics into a general-purpose BIQA method named blind PRI-based (BPRI) metric. The BPRI metric is opinion-unaware (OU) and almost training-free except for the distortion identification process. Comparative studies on five large IQA databases show that the proposed BPRI model is comparable to the state-of-the-art opinion-aware- and OU-BIQA models. Furthermore, BPRI not only performs well on natural scene images, but also is applicable to screen content images. The MATLAB source code of BPRI and other PRI-based distortion-specific metrics will be publicly available.**

***Index Terms*—Blind image quality assessment (BIQA), pseudo-reference image (PRI), blockiness, sharpness, noisiness.**

Manuscript received September 12, 2017; accepted November 29, 2017. Date of publication December 29, 2017; date of current version July 17, 2018. This work was supported by the National Natural Science Foundation of China under Grants 61422112, 61371146, 61521062, and 61527804. Part of this work was presented at the 2016 IEEE International Conference on Multimedia and Expo [1]. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Wolfgang Hürst. *(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).

J. Liu is with the School of Electrical and Information Engineering, Tianjin University, Tianjin 300072, China (e-mail: jliu_tju@tju.edu.cn).

C. W. Chen is with the State University of New York at Buffalo, Buffalo, NY 14260 USA (e-mail: chencw@buffalo.edu).

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/TMM.2017.2788206

I. Introduction

THE ubiquitous degradations existed in various visual communication systems call for efficient and effective image quality assessment (IQA) metrics which can predict the perceptual quality of images accurately. The ultimate goal of objective IQA is to develop computational metrics which can compute quality scores well correlated to human perceptions. During the past decades, dozens of IQA models have been proposed [2]–[4]. Although there are many distinctive ways to solve this problem, most of them share a common framework. In visual communication systems, the target image is usually degraded from an original perfect quality image called reference image. According to the availability of the reference image, IQA metrics can be classified into full-reference (FR), reduced-reference (RR) and no-reference (NR).

FR IQA metric can be interpreted as an image fidelity metric which describes the similarity between two images. It usually measures the target image's deviation from the perfect quality reference image [2], [3]. If we describe the problem in a "quality axis", FR IQA metric measures the "distance" from the reference image to the target image. Although sometimes the reference image is not available, we can still derive some characteristics of the perfect quality images, and then measure the distorted image's deviation from such characteristics [4]. Therefore, whether the reference image is given or not, a common spirit of current IQA metrics is that: the perfect quality image is considered as the "reference point" of those methods, and the quality of the distorted image can be measured as a "distance" from this reference point.

Under such a framework, naturally a question can be asked is that "what about the opposite direction of the quality axis?" Few literature has discussed this question. If we can define an extreme point in the opposite direction of the quality axis which indicates the worst quality, then we can take it as a new "base point" for quality assessment. Based on this motivation, we generate a pseudo-reference image (PRI) in this work. Fig. 1 illustrates a comparison of PRI-based quality metric and traditional IQA metric. Contrary to the conventional perfect quality reference image, the PRI is located in the opposite direction of the quality axis. It is derived from the distorted image, and it suffers from the severest distortion. After the generation of PRI, we can take the PRI as a new base point, and measure the distorted image's distance from the PRI as the quality. Crete *et al.* [5] and Li *et al.* [6] proposed blur metrics by comparing a blurred image and the re-blurred image. The idea is close to the PRI-based

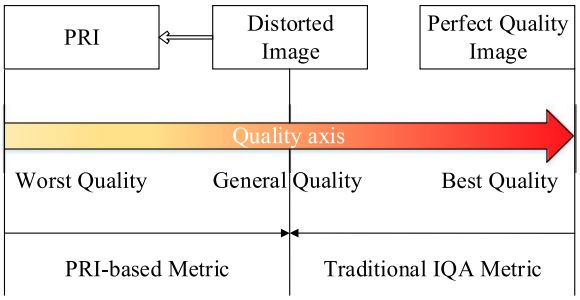


Fig. 1. A comparison of PRI-based quality metric and traditional IQA metric.

IQA, but the proposed metrics can only estimate sharpness. In this work, we introduce the concept of PRI, and discuss the PRI-based IQA framework systematically. We demonstrate that the PRI-based framework is a general IQA framework, which works not only for one specific distortion but also for various other distortions. Moreover, the PRI-based framework can be also used for general-purpose IQA by integrating PRI-based distortion-specific IQA metrics.

To develop PRI-based IQA metrics, the following problems need to be solved: (1) the definition of PRI; (2) the measure of the quality distance. Since each type of distortion causes specific artifacts, we need to define distortion-specific PRI to be consistent with the characteristics of a given distortion. Also it requires different distance measures for different distortion types. In this work, we focus on blocking, blurring and noising artifacts since these are the most common artifacts in practical visual communication systems, and many other complex distortions can be combinations of these artifacts. We develop PRI-based quality metrics to estimate blockiness, sharpness and noisiness. The PRI-based distortion-specific metrics are then integrated into a general-purpose blind IQA (BIQA) method through a 2-stage quality regression after distortion identification framework, which is used in DIIVINE [7] and performs distortion-specific quality assessment after distortion identification. Fig. 2 illustrates a framework of the proposed PRI-based general-purpose BIQA model. It is worth mentioning that the proposed PRI-based method is opinion-unaware (OU), and almost training-free except for the distortion identification process.

Block-based compression causes blocking artifacts because of the independent quantization of each individual block. The blocking artifacts at the block boundaries together can form some pseudo structures. We find that such pseudo structures show certain degree of similarity in images compressed at various compression levels. The pseudo structure and the genuine image content structure both exist in the distorted image, but they can be distinguished according to their positions. The genuine image content structures can be everywhere, while pseudo structures are distributed at the block boundaries. To obtain the PRI, we compress the distorted image to the utmost using the highest compression level in the encoder. The distorted image's and the PRI's pseudo structures become more and more similar as the compression level gets higher and higher. In this paper, we propose a PRI-based blockiness metric named pseudo structure similarity (PSS) by calculating the similarity between pseudo structures of the target image and the corresponding PRI.

Blurring and noising distortions both degrade the textures of images, which significantly change the local image structures. For example, some texture patterns such as edges or corners will be turned into flat patterns because of blurring; whereas flat patterns can be turned into texture patterns because of noising. Thus measuring the change of local structures can be an effective way to measure both sharpness and noisiness, and we can utilize the same framework to estimate both artifacts. However, it is not easy to measure sharpness/noisiness using sole blurred/noisy images since local structures are highly sensitive to image content. In this paper, we compare the blurred/noisy images with the PRIs to reduce the influence of image content. To get the PRI, we further blur the current blurred image using a specific smoothing filter, and add a certain intensity of noise to the current noisy image. The further degraded images are taken as the PRIs. Local binary pattern (LBP) [8] is selected to describe the local structures because of its simplicity and high efficiency. We calculate the similarity of specific LBPs between the blurred/noisy image and the corresponding PRIs. The proposed PRI-based sharpness/noisiness metrics are referred as local structure similarity (LSS), specifically LSS_s for sharpness estimation and LSS_n for noisiness estimation. More blurred/noisy images have higher LSS scores since blurring/noising can only cause significant LBP changes in sharp/clean images.

The proposed PSS and LSS metrics both follow the framework of PRI-based IQA illustrated in Fig. 1. We solve the problem of IQA by measuring "how much worse the image could be?" rather than the traditional "how bad the image is?". As illustrated in Fig. 2, we then develop a general-purpose blind PRI-based (BPRI) IQA model by integrating the distortion-specific PRI-based IQA metrics through a 2-stage framework, in which we identify distortions first and then perform distortion-specific quality assessment. We compare the BPRI model with the state-of-the-art opinion-aware (OA) and opinion-unaware (OU) general-purpose BIQA metrics on 5 large IQA databases, including LIVE [9], TID2013 [10], CSIQ [11], SIQAD [12] and CCT [13]. The SIQAD database is a recently-constructed screen content image (SCI) quality assessment (QA) database, and the CCT database focuses on cross-content-type IQA. Experimental results show that PRI-based IQA is effective, and the proposed BPRI model is superior or comparable to the state-of-the-art OA- and OU-BIQA metrics.

The remainder of this paper is organized as follows. In Section II, we review some related works, including the state-of-the-art distortion-specific and general-purpose BIQA models. In Section III, we describe the details of the proposed PRI-based distortion-specific metrics, i.e., PSS and LSS. In Section IV, we present the proposed general-purpose BPRI model. The experimental results are given in Section V. Finally, Section VI concludes this paper and gives some discussions.

II. Related Work

A. Distortion-Specific BIQA Metric

1) Blockiness Estimation: Blockiness is a common type of distortion easily encountered in the visual communication systems. It is usually caused by the relatively independent processing of individual blocks which is involved in most block-based

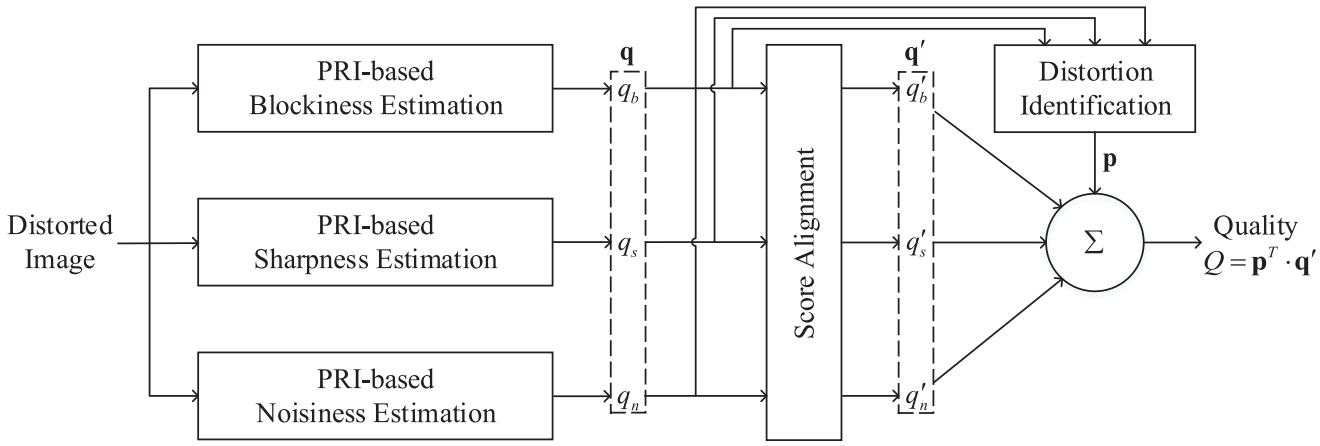


Fig. 2. A framework of the proposed PRI-based general-purpose BIQA metric, which is integrated from the proposed PRI-based distortion-specific metrics through a 2-stage quality regression after distortion identification framework.

compression schemes. During recent years, many IQA models have been proposed to evaluate the perceptual quality of JPEG images [1], [14]–[19]. Intuitively, JPEG image can be evaluated spatially from the inter-block blocking effect and intra-block blurring effect because of the within-block quantization and the discard of high frequency information. Liu and Heynderickx [16] combined the local blocking distortion with its local visibility through visual masking. The regularities of pseudo structures were measured in [18]. Zhan and Zhang [19] considered blockiness and the luminance change within blocks. Wang *et al.* [15] estimated blockiness through pixel changes across block boundaries, and also estimated blurring through both within-block differences and the zero-crossing rate of the differencing image. In a previous work [1], we proposed a pseudo structural similarity metric based on the PRI-based IQA framework.

The problem can be also considered from the transform domain. Bovik and Liu [14] modeled the blocking artifact as a 2-D step function by constructing a new block from two neighbouring blocks. They constructed the new block and extracted needed parameters in the discrete cosine transform (DCT) domain. In [17], the number of zero-valued DCT coefficients was counted in each block, and the counts were then weighted through a quality relevance map.

2) Sharpness Estimation: Sharpness estimation is also a nontrivial work because of the wide spread of blurring distortion in various application scenes such as picture capturing and image/video coding. Early sharpness estimators try to measure sharpness via edge analysis. Ferzli and Karam [20] introduced a concept of just noticeable blur (JNB), which described the minimum blurring needed to be perceived near the edge given a contrast higher than the just noticeable difference (JND). Based on JNB, Narvekar and Karam [21] developed a probabilistic model considering human's sensitivity of blur at various contrasts. Li *et al.* [22] proposed a sparse-based sharpness metric since the spread of edges could be captured by the sparse coefficients.

Estimating sharpness from the transform domain is also a good strategy, since blurred image is often short of high-frequency information. Vu and Chandler [23] applied a three-level separable discrete wavelet transform (DWT) and then a weighted average of the log-energies of the DWT subbands was computed as the sharpness. A spectral and spatial sharpness (S_3) was proposed in [24]. S_3 combined a spectral estimator with a spatial estimator. Since blurring introduced losses of local phase coherence (LPC), Hassen *et al.* [25] proposed a LPC-based sharpness metric. Gu *et al.* [26] estimated sharpness in an autoregressive parameter space.

3) Noisiness Estimation: Noise estimation has been an important and fundamental problem in the field of image processing and computer vision. This problem can be solved by estimating the intensity (typically the variance) of noise [27]–[29]. Zoran and Weiss [27] estimated noise based on the assumption that clean image should have a constant kurtosis value throughout scales, and deviations from this were due to noise. Liu *et al.* [28] inferred the noise as a function of image intensity from a single image. Tang *et al.* [29] estimated noise through statistical analysis and noise injection. However, the intensity of noise sometimes can be deviated from the perceptual quality of the noisy image. Thus some researchers solve this problem from an IQA perspective of view, and propose noise-specific IQA metrics [30], [31]. Zhai *et al.* [30] proposed a dual-model approach based on the assumption that human visual system had different behavioral patterns under low and high noise levels. Liu *et al.* [31] proposed a blind IQA model for noisy image based on the free energy principle [32].

B. General-Purpose BIQA Metric

1) OA Metric: Some general-purpose BIQA models are based on natural scene statistics (NSS). DIIVINE [7] identifies distortion first, and then conducts distortion-specific IQA using NSS of the wavelet coefficients. BRISQUE [33] uses the scene statistics of local luminance coefficients after mean subtraction and divisive normalization to quantify possible losses of "naturalness". BLIINDS-II [34] utilizes NSS of DCT coefficients. Some other measures integrate handcrafted features which also rely on the statistical regularities of natural images. For example, the free energy principle is utilized in [32] and [35]. Learning-based features can be utilized for IQA. For example, unsupervised feature learning is used in CORNIA [36]

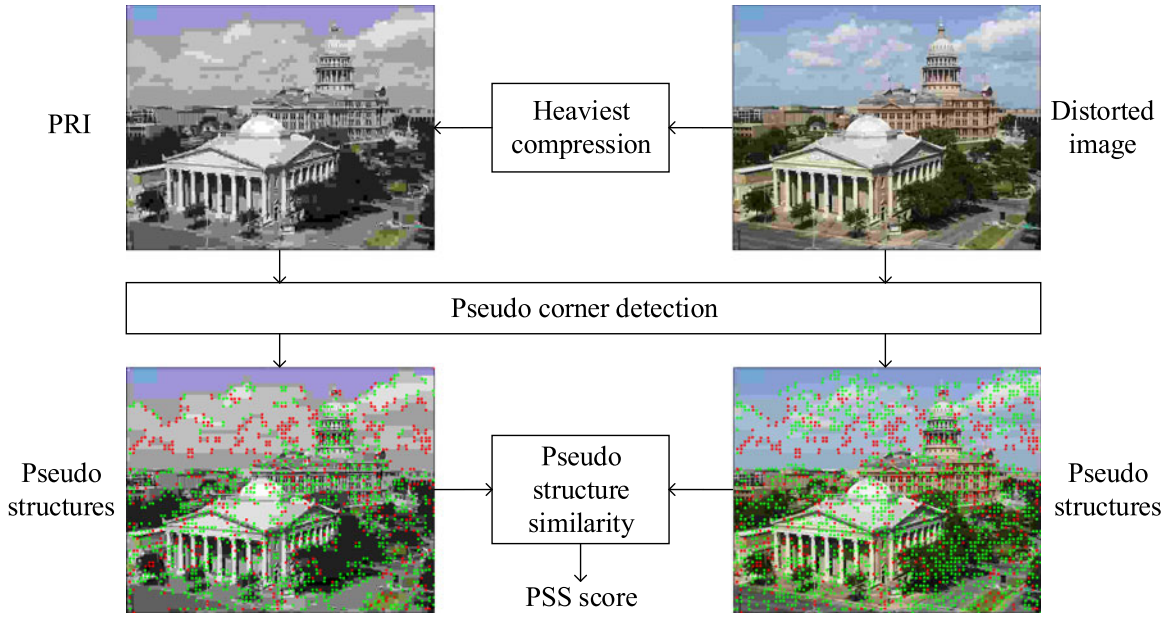


Fig. 3. A framework of the PRI-based blockiness metric PSS. The red and green dots in the bottom 2 images denote the pseudo corners. Specially, the red ones indicate that pseudo corners in the distorted image and PRI are overlapped. All pseudo corners can give an overall description of the pseudo structures.

and HOSA [37]. Li *et al.* [38] considered statistical structural and luminance features. Although different strategies are considered in the above metrics, a common thing is that they are all OA metrics. Ground-truth quality scores and support vector regression are needed to learn the mapping from the quality features to the final quality score.

2) OU Metric: Due to the limited generalizability of the OA metrics, some OU metrics which use no opinion scores for training are proposed. NIQE [39] utilizes spatial domain NSS. It uses multivariate gaussian models (MVGs) to fit the NSS features, and the quality is measured as the distance between MVGs. IL-NIQE [40] follows the same framework, but different features are used. QAC [41] performs quality-aware clustering and learns a set of quality-aware centroids to act a codebook to infer the quality. LPSI [42] is a training-free method which utilizes statistical features extracted from binary patterns of local image structures. OU metrics have strong generalizability, but limited metrics are available now.

III. PRI-Based Distortion-Specific Metrics - PSS and LSS

A. PRI-Based Blockiness Measure - PSS

As described in Section I, the proposed blockiness metric PSS is based on the PRI. We first compress the distorted image to the PRI using the heaviest compression level in the encoder. The pseudo structures of the distorted image and the corresponding PRI are then detected, and finally the blockiness is measured as the similarity between two images' pseudo structures. Fig. 3 illustrates a framework of the PSS metric. More details are as follows.

1) The PRI: The PRI has a worse quality than the distorted image, and it is compressed from the distorted image. Given the distorted image $\mathbf{A}$, we derive the PRI $\mathbf{M}$ by

$$\mathbf{M} = \mathrm{JPEG}(\mathbf{A}, QT), \tag{1}$$

where JPEG denotes the JPEG encoder, and QT is the utilized quantization table which is fixed and represents a very low compression quality. Although we use the PRI as a "reference", PSS is still a NR model since the PRI is derived from the distorted image. As shown in Fig. 3, the upper left image is the PRI of the upper right image.

2) Pseudo Corners and Structures: Since corners can describe image structures, they are frequently used in various applications such as video tracking and motion detection [43]. Corners could be also used to estimate blockiness, for example, Min *et al.* [13] considered the regularity of corners by quantifying the ratio of pseudo corners. In this paper, we also detect corners, and then use them to represent image structures. In JPEG compressed images, image structures are composed of both genuine image content structure and artificial pseudo structure resulted from excessive compression. We differentiate them by analysing the positions of the detected corners. If the detected corners are distributed at 4 corners of the 8×8 block, they are identified as pseudo corners. Otherwise, if they are detected at some ordinary positions, they are taken as regular corners. Then the pseudo structures of an image can be described by all detected pseudo corners.

Given an image $\mathbf{A} = (a_{ij})_{h\times w}$, its pseudo structure is defined as $\mathbf{P} = (p_{ij})_{h\times w}$, where h, w indicate the image rows and columns, and the elements are

$$p_{ij} = \begin{cases} 1 & \text{if } a_{ij} \in C, \mathrm{mod}(i,N) < 2, \mathrm{mod}(j,N) < 2 \\ 0 & \text{otherwise} \end{cases}, \tag{2}$$

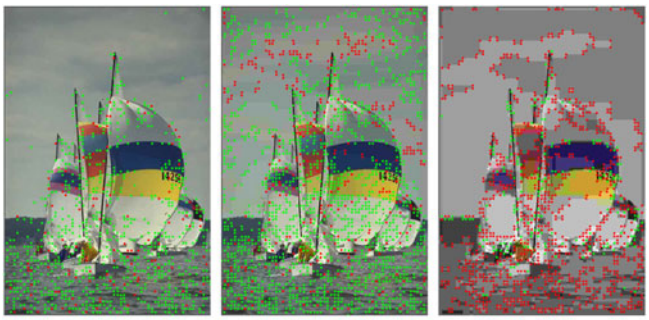


Fig. 4. Pseudo structures in images of different compression degrees. DMOS is the difference mean opinion score, and PSS is the quality score computed by the PSS metric. Similar to Fig. 3, the red and green dots denote the pseudo corners. There are more overlapping pseudo corners in more distorted images. (a) DMOS = 22, PSS = 0.09. (b) DMOS = 42, PSS = 0.26. (c) DMOS = 90, PSS = 0.82.

where $a_{ij} \in C$ means that a corner is detected at pixel a_{ij}; mod calculates the remainder; $\mathrm{mod}(i, N) < 2$, $\mathrm{mod}(j, N) < 2$ denotes position (i, j) is one of the block corner, and the block size is $N = 8$. Shi and Tomasi's minimum eigenvalue method [43] is used to detect corners. We use $\mathbf{P}_d = (p_{d\,ij})_{h\times w}$ and $\mathbf{P}_m = (p_{m\,ij})_{h\times w}$ to denote the pseudo structures of the distorted image and the PRI. As illustrated in the bottom two images of Fig. 3, the red and green dots denote pseudo corners ($p_{ij} = 1$). All dots together give a description of pseudo structures $\mathbf{P}_d$ and $\mathbf{P}_m$.

3) Pseudo Structure Similarity: As illustrated in Fig. 3, if we compare $\mathbf{P}_d$ and $\mathbf{P}_m$, we can find certain degree of similarity between them, and there are lots of overlapping pseudo corners. We use $\mathbf{P}_o = (p_{o\,ij})_{h\times w}$ to denote the overlapping pseudo structures, i.e., the overlap between $\mathbf{P}_d = (p_{d\,ij})_{h\times w}$ and $\mathbf{P}_m = (p_{m\,ij})_{h\times w}$

$$\mathbf{P}_o = (p_{o\,ij})_{h\times w} = (p_{d\,ij} \cdot p_{m\,ij})_{h\times w}. \tag{3}$$

The red dots in Figs. 3 and 4 describes $\mathbf{P}_o$. Fig. 4 illustrates pseudo structures in images of different compression degrees. Note that when the image is more compressed, its pseudo structures $\mathbf{P}_d$ will be more similar to the PRI's pseudo structures $\mathbf{P}_m$, i.e., there are more overlapping pseudo corners.

Thus we introduce pseudo structure similarity (PSS) to estimate blockiness. We define the number of overlapping pseudo corners as N_o, and the number of pseudo corners in $\mathbf{P}_m$ as N_m

$$N_o = \sum_{i,j} p_{o\,ij}\,; \quad N_m = \sum_{i,j} p_{m\,ij}. \tag{4}$$

Then the pseudo structure similarity (PSS) is described as

$$\mathrm{PSS} = \frac{N_o}{N_m + 1}, \tag{5}$$

where the constant 1 is added for stability reason. Intuitively, PSS score denotes the overlapping degree between $\mathbf{P}_d$ and $\mathbf{P}_m$. Severer blockiness results in higher PSS value. As shown in Fig. 4, PSS correlates well with the blockiness degree and perceptual quality. More precise quantitative validation of the proposed method is given in Section V.

4) Implementation Details: Note that PSS involves very few parameters. We only need to set several parameters when deriving the PRI and detecting corners. In (1), we use the JPEG encoder in MATLAB. Specifically, the "imwrite" function is used. The quantization table QT corresponds to the "Quality" parameter of "imwrite", which is set to the extreme point 0. It indicates the severest compression the encoder can provide. In (2), we use the MATLAB implementation of the minimum eigenvalue method [43] to detect corners. Specifically, "corner" function is used. The maximum number of detected corners is set as a very large number, which means that we do not bound the number in normal situations. A 3×3 Gaussian mask with a standard deviation of 0.5 is used to filter the target image first. The "QualityLevel" parameter which specifies the minimum corner quality is set as 0.001. The quality parameter in PSS is set as a very small value so that we can detect more corners to describe the structures.

B. PRI-Based Sharpness and Noisiness Metrics - LSS

Following the same PRI-based IQA framework described above, we propose PRI-based metrics LSS_s and LSS_n to estimate sharpness and noisiness. Fig. 5 illustrates a framework of LSS. LSS_s and LSS_n both follow this framework. The differences are that LSS_s and LSS_n have different PRI definitions, and different LBPs are extracted to describe the local structures. In LSS_s, the current blurred image is blurred to the PRI using a specific smoothing filter; while in LSS_n, certain intensity of noise is added to the current noisy image to get the PRI. Different LBPs are then extracted to describe the local structures, since different LBPs can be sensitive to different distortions. Finally, LSS calculates the similarity between LBPs extracted from the distorted image and the PRI as the quality score.

1) The PRI: The PRI in LSS_s is blurred from the given distorted image $\mathbf{A}$

$$\mathbf{M} = \mathbf{f} \otimes \mathbf{A} = \frac{1}{9}\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \otimes \mathbf{A}, \tag{6}$$

where $\mathbf{M}$ is the derived PRI; $\mathbf{f}$ is the blurring filter, and we set it as an averaging filter; $\otimes$ is the convolution operator.

The PRI in LSS_n is created by adding Gaussian noise to the given distorted image $\mathbf{A}$

$$\mathbf{M} = \mathbf{A} + \mathcal{N}(0, v), \tag{7}$$

where $\mathbf{M}$ is the created PRI; $\mathcal{N}(0, v)$ generates normally distributed random values with 0 mean and v variance. Similar to PSS, LSS_s and LSS_n are also NR metrics. Fig. 5 shows an example of PRI in LSS_s.

2) Local Binary Patterns and Structures: LBP [8] is a powerful visual descriptor frequently used in many image processing and computer vision applications. LBP was also previously used in IQA [38], [42], [44]. Wu *et al.* [42] utilized statistical features of local image structures described by LBPs. Li *et al.* [38] considered the histograms of LBPs as one part of the integrated quality features. Wu *et al.* [44] proposed a RR IQA metric by measuring the change of structural histogram described through

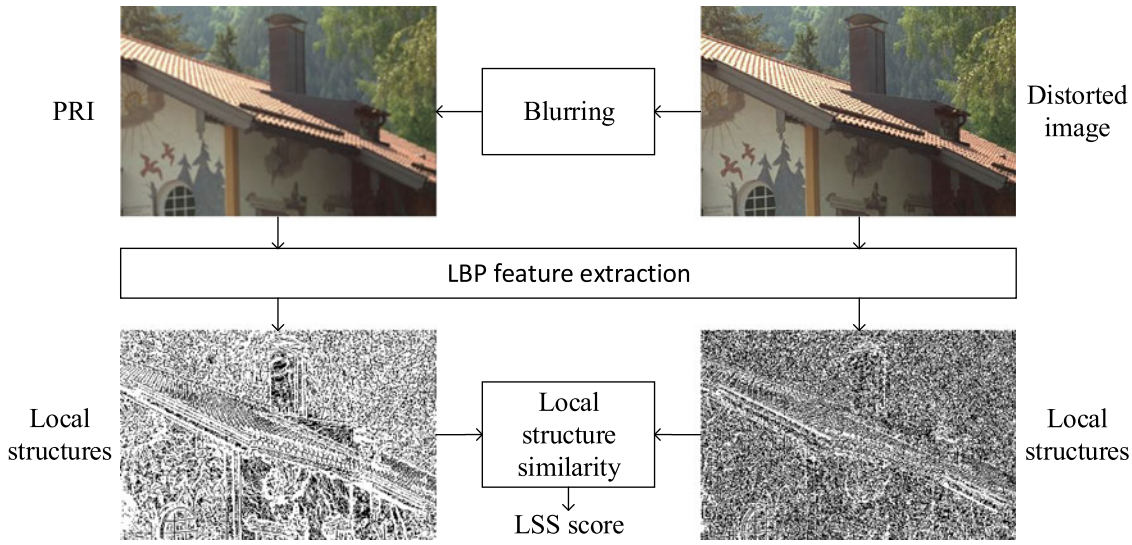


Fig. 5. A framework of the PRI-based metric LSS (we illustrate LSS_s as an example here). The white pixels in the bottom two maps denote the desired local structures are detected. Local structures are described by LBPs, and LSS calculates the similarity between particular LBPs of the distorted image and PRI.

LBP. The underlying assumption of the above LBP-based IQA metrics is that quality degradation can change local image structures which can be modeled by LBPs. For example, blurring can smooth the textures of images, and turn some texture patterns into flat patterns; whereas flat patterns can be turned into texture patterns because of noising. So we extract LBPs and quantify such local structure change to measure sharpness and noisiness.

LBP compares the luminance values of a center pixel g_c and its circularly symmetric neighborhoods g_p, and then binarizes the differences between them using an unit step function, and codes the binarization results by summing them into a numerical value

$$\mathrm{LBP}_{\mathrm{P,R}} = \sum_{p=0}^{P-1} u(g_p - g_c), \tag{8}$$

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

$$u(x) = \begin{cases} 1 & x \geq 0 \\ 0 & x < 0 \end{cases}. \tag{9}$$

Note that the LBP definition in (8) is different from the non-uniform definition in [8], where the authors code the binarization results by attaching a factor 2^p to the binary results of pixel g_p. It is also slightly different from their uniform definition. Here we do not code the LBP which has many spatial transitions as a separate number for simplicity.

In this paper, we set $P = 4$ and $R = 1$ for simplicity. It does not require any interpolation and the possible LBP values are also quite few. We find that several specific LBPs are sensitive to blurring, while some other LBPs are sensitive to noising. From the blurred image to the corresponding PRI, $\mathrm{LBP}_{4,1} = 2$ and $\mathrm{LBP}_{4,1} = 3$ change the most significantly; while from the noisy image to the corresponding PRI, $\mathrm{LBP}_{4,1} = 0$ and $\mathrm{LBP}_{4,1} = 1$ change the most significantly. Thus we quantify the changes of different LBPs as the quality metric for blur and noise. In LSS_s, we define the desired local structure map as $\mathbf{L} = (l_{ij})_{h\times w}$, where the elements are

$$l_{ij} = \begin{cases} 1 & \text{if } \mathrm{LBP}_{4,1} = 2 \text{ or } \mathrm{LBP}_{4,1} = 3 \\ 0 & \text{otherwise} \end{cases}. \tag{10}$$

Similarly, the local structure map $\mathbf{L} = (l_{ij})_{h\times w}$ in LSS_n is defined as

$$l_{ij} = \begin{cases} 1 & \text{if } \mathrm{LBP}_{4,1} = 0 \text{ or } \mathrm{LBP}_{4,1} = 1 \\ 0 & \text{otherwise} \end{cases}. \tag{11}$$

We adopt the same calculation process for both images, and denote the local structure maps of the distorted image and the PRI as $\mathbf{L}_d = (l_{d\,ij})_{h\times w}$ and $\mathbf{L}_m = (l_{m\,ij})_{h\times w}$, respectively. The bottom two images in Fig. 5 illustrate examples of $\mathbf{L}_d$ and $\mathbf{L}_m$.

3) Local Structure Similarity: As described above, we measure the similarity between $\mathbf{L}_d = (l_{d\,ij})_{h\times w}$ and $\mathbf{L}_m = (l_{m\,ij})_{h\times w}$ as the quality. We define the overlap between them as $\mathbf{L}_o = (l_{o\,ij})_{h\times w}$, which can be calculated as

$$\mathbf{L}_o = (l_{o\,ij})_{h\times w} = (l_{d\,ij} \cdot l_{m\,ij})_{h\times w}. \tag{12}$$

Fig. 6 illustrates some $\mathbf{L}_o$ examples of images of different blurring and noising degrees. Similar to PSS, there are more overlap between the feature maps of the distorted image and the PRI in more distorted images. We can also define the union between $\mathbf{L}_d$ and $\mathbf{L}_m$ as $\mathbf{L}_u = (l_{u\,ij})_{h\times w}$, whose elements are

$$\mathbf{L}_u = (l_{u\,ij})_{h\times w} = (l_{d\,ij} | l_{m\,ij})_{h\times w}. \tag{13}$$

Then the proposed local structure similarity (both LSS_s and LSS_n) can be defined as

$$\mathrm{LSS} = \frac{N_o}{N_u + 1}, \tag{14}$$

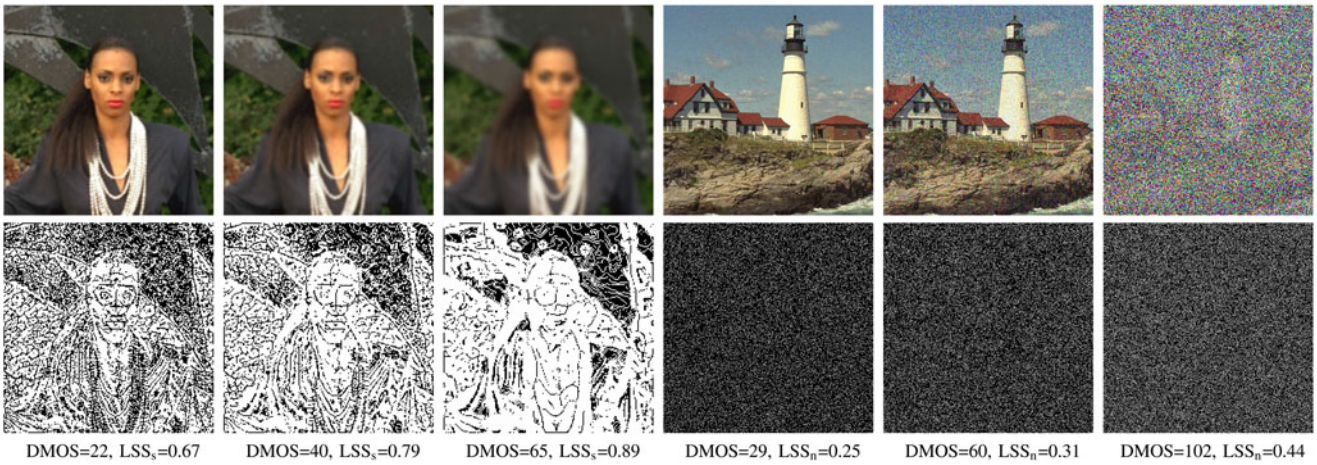


Fig. 6. Local structure similarity in images of different distortion levels. 1st row: cropped distorted images; 2nd row: the corresponding $\mathbf{L}_o$ maps, i.e., the overlap between the distorted image's and the PRI's local structure maps; DMOS: difference mean opinion score; LSS$_s$, LSS$_n$: the quality scores computed by the LSS metrics. There are more overlap in more distorted images.

where N_o and N_u mean the number of non-zero elements in the $\mathbf{L}_o$ and $\mathbf{L}_u$ maps, which can be calculated through

$$N_o = \sum_{i,j} l_o{}_{ij}\,; \quad N_u = \sum_{i,j} l_u{}_{ij}. \tag{15}$$

Similar to PSS, LSS indicates the overlapping degree between $\mathbf{L}_d$ and $\mathbf{L}_m$, and more blurred/noisy images have higher LSS values. As shown in Fig. 6, LSS scores provide good predictions of the sharpness and noisiness degrees. Note that LSS adopts a global averaging strategy. Considering the success of percentile pooling in sharpness estimation, LSS may be improved by incorporating visual attention [45]–[47].

4) Implementation Details: Compared with PSS, LSS$_s$ and LSS$_n$ involve even fewer parameters. As described above, LSS$_s$ uses a 3×3 averaging filter to derive the PRI, and Gaussian noise with 0 mean and 0.5 variance is added to create the PRI in LSS$_n$. In both LSS metrics, $\mathrm{LBP}_{4,1}$ is used to calculate the local structure maps.

IV. PRI-Based General-Purpose Metric - BPRI

Motivated by DIIVINE [7], we integrate the above PRI-based distortion-specific metrics, i.e., PSS, LSS$_s$ and LSS$_n$, into a general-purpose blind PRI-based (BPRI) metric through a 2-stage quality regression after distortion identification framework. Different from DIIVINE [7] which needs to learn a regressor to integrate a large group of features, single quality scores computed by PSS, LSS$_s$ and LSS$_n$ are effective enough to describe the quality degradation caused by blockiness, blur and noise. Thus it is feasible to integrate PSS, LSS$_s$ and LSS$_n$ in an opinion-unaware manner.

A. Score Alignment

First we need to align the quality scores, since PSS, LSS$_s$ and LSS$_n$ respond differently to the quality variations. The quality scores computed by PSS, LSS$_s$ and LSS$_n$ are denoted as q_b, q_s and q_n, respectively. We use the following five-parameter logistic function to map the quality scores

$$q' = \lambda_1 \left(\frac{1}{2} - \frac{1}{1 + e^{\lambda_2 (q - \lambda_3)}} \right) + \lambda_4 q + \lambda_5, \tag{16}$$

where q, q' are the original and mapped quality scores, respectively; $\{\lambda_i | i = 1, 2, \ldots, 5\}$ are five parameters determined through curve fitting. Quality scores mapped from q_b, q_s and q_n are denoted as q'_b, q'_s and q'_n, respectively.

To fit the parameters, we create an image set having no content overlap with the widely used IQA databases. Specifically, we select 100 high quality natural images from a public database [48], and generate the distorted images using the 4 most common types of distortions, i.e., JPEG compression (JPEG), Gaussian blur (GB), white Gaussian noise (WN), and JPEG2000 compression (JP2K). For each distortion type, 5 quality levels are selected to cover a wide range of perceptual quality. A total of 2,000 distorted images are used to fit the parameters. Target scores are also needed to conduct the curve fitting. In IQA model evaluation, the ground-truth quality scores are used. In consideration of the successes of FR IQA metrics, we use FR metric to guide the score alignment, i.e., all distortion-specific quality scores are non-linearly mapped to the quality scores of FR metric. GMSD [49] is used in BPRI. We will also try other FR metrics, and test the performance in Section V. After score alignment, q'_b, q'_s and q'_n are comparable and respond consistently to the quality variations.

B. Distortion Identification and Quality Regression

We train a classifier to identify the distortions. For simplicity, we utilize q_b, q_s and q_n as features for the identification. The image set described in Section IV-A is used as the training set Φ. Considering the successes of support vector machine (SVM) in classification problems, we select SVM for classification. Given the features $\mathbf{q} = [q_b, q_s, q_n]^T$ and the label $l \in \{b, s, n\}$ of each image and the training set Φ, we can train the $classifier$

$$classifier = \mathrm{SVM\_TRAIN}(\mathbf{q}_i, l_i), i \in \Phi, \tag{17}$$

where i is the image index. Then given the quality feature $\mathbf{q}$ of any test image, we can identify distortions using the pre-trained $classifier$

$$\mathbf{p} = \text{SVM\_PREDICT}(\mathbf{q}, classifier), \tag{18}$$

where $\mathbf{p} = [p_b, p_s, p_n]^T$ are the predicted probabilities of different distortions.

LIBSVM [50] is adopted to implement SVM with a radial basis function (RBF) kernel. We follow the common SVM parameter settings used in the training of OA-BIQA metrics, and only change the model type from "epsilon-SVR" to "C-SVC", i.e., from regression to classification. During the score alignment and classifier training, the JP2K subset is labeled as GB since we do not design distortion-specific metric for JP2K and the main distortion in JP2K compressed images is blur. We adopt a probability weighting strategy and the final quality score of BPRI is then described as

$$Q = \mathbf{p}^T \cdot \mathbf{q}', \tag{19}$$

where $\mathbf{q}' = [q'_b, q'_s, q'_n]^T$ are the aligned scores. As an alternative, we can also adopt a hard classification strategy, in which we classify the distortion type explicitly and use the corresponding aligned score as the quality. We will also test the performance of such strategy in Section V.

V. Experimental Results

To validate the proposed BPRI quality model, first we test BPRI and the state-of-the-art BIQA models on the four most common distortion types mentioned above, then we will test all BIQA models' generalizability to other distortions. We test BIQA models on images of two most dominant types we may encounter in realistic visual communication systems, i.e., natural scene image (NSI) and screen content image (SCI). SCI is also an indispensable part of visual communication systems, and SCI QA becomes more and more important because of the widespread of the so-called "screen content" which is often computer-generated and thus has some extraordinary characteristics [12], [13].

A. Experimental Protocol

1) Databases: Five large IQA databases are used as testbeds, including LIVE [9], TID2013 [10], CSIQ [11], SIQAD [12] and CCT [13]. The first 3 are commonly used NSI QA databases. The SIQAD database is constructed to facilitate the research of SCI QA [12]. The CCT is a cross-content-type IQA database constructed to test IQA models' generalizability to other content types [13]. It includes 1,320 NSIs, computer graphic images (CGIs) and SCIs compressed by the high efficiency video coding (HEVC) intra coding method and the screen content compression (SCC) extension of HEVC. The main artifacts introduced by HEVC and HEVC-SCC are blurring and blockiness. Current IQA studies are mostly concentrated on NSI, and there are more publicly available NSI QA databases. The first four databases share four common distortion types, i.e., JPEG, GB, WN and JP2K, and there are 634, 500, 600 and 560 commonly distorted images in each database, respectively. We will test BIQA models on the commonly distorted images first, and then test their distortion generalizability on other non-common distortions included in the LIVE, TID2013, CSIQ, SIQAD, and CCT databases.

2) Comparing Algorithms: We compare the proposed BPRI model with the state-of-the-art BIQA models, including:

- OA-BIQA models: DIIVINE [7], BLIINDS-II [34] and BRISQUE [33] use wavelet, DCT and spatial domain NSS, respectively. NFERM [35] is based on free energy principle. CORNIA [36] and HOSA [37] are both based on unsupervised feature learning.
- OU-BIQA models: NIQE [39] and IL-NIQE [40] use spatial domain NSS. QAC [41] performs quality-aware clustering and learns a codebook. LPSI [42] utilizes local image structure statistics.

We use the original implementations released by the authors. A short review of the above BIQA models can be found in Section II-B.

3) Evaluation Criteria: To evaluate the IQA models, we follow the common procedures [13], [51], [52], and first map the predicted quality scores using the five-parameter logistic function as shown in (16). Different from the score alignment, q and q' now are the predicted and mapped quality scores; $\{\lambda_i | i = 1, 2, \ldots, 5\}$ are fitted parameters using the predicted quality scores and the ground-truth quality scores. Then the consistency between the mapped score and the ground-truth is calculated to measure the performance of the IQA model. We choose the following 3 mainstream metrics as the evaluation criteria.

- Spearman rank-order correlation coefficient (SRCC). It denotes the monotonicity between the ground truth and predicted score.
- Pearson linear correlation coefficient (PLCC), which measures the IQA metric's prediction linearity.
- Root-mean-square error (RMSE), which calculates the error between the mapped score and the ground truth.

B. Performance on Common Distortions

We first test all BIQA models on common distortions. Table I lists the performance comparison results. Besides single distortions, we also test on all 4 types of distorted images. All OA-BIQA models are trained on the LIVE, thus we do not list their performance on the LIVE in Table I to ensure complete separation of training and testing. We mainly compare the BPRI with OU-BIQA models on the LIVE. Similarly, CORNIA and HOSA learn a codebook using images from the CSIQ, thus their performance is excluded on the CSIQ. For the proposed BPRI method, both the probability weighting strategy and the hard classification strategy are tested, which are denoted as BPRI(p) and BPRI(c), respectively. We have highlighted the top 2 models according to each criterion. Mean SRCC and PLCC performance averaged over all databases and the total hit count (the number of times ranked in the top 2) are also reported in Table I.

It can be observed that the proposed BPRI models remain on the top on most databases, especially on the SIQAD where the proposed methods show remarkable superiority. Most of the current models are not so good at handling SCIs, while the proposed

TABLE I
PERFORMANCE COMPARISON ON COMMON DISTORTIONS

Criteria	Database	Dist.	OA Models						OU Models					
			DIIVINE	BLIINDS-II	BRISQUE	NFERM	CORNIA	HOSA	NIQE	QAC	IL-NIQE	LPSI	BPRI(c)	BPRI(p)
SRCC	LIVE	JPEG	–	–	–	–	–	–	0.9410	0.9362	0.9424	**0.9677**	0.9665	**0.9699**
		GB	–	–	–	–	–	–	**0.9326**	0.9134	0.9154	0.9156	**0.9268**	0.9243
		WN	–	–	–	–	–	–	0.9716	0.9509	0.9809	0.9557	**0.9843**	**0.9854**
		JP2K	–	–	–	–	–	–	**0.9187**	0.8621	0.8944	**0.9300**	0.9080	0.9069
		All	–	–	–	–	–	–	0.9168	0.8857	0.9153	0.8333	**0.9288**	**0.9304**
	TID2013	JPEG	0.6288	0.8360	0.8448	0.8722	0.8958	0.8957	0.8468	0.8369	0.8340	**0.9123**	0.9067	**0.9107**
		GB	0.8344	0.8367	0.8137	0.8501	**0.9274**	0.8695	0.7986	0.8464	0.8148	0.8408	**0.8725**	0.8593
		WN	0.8553	0.6468	0.8520	0.8581	0.7354	0.8172	0.8187	0.7427	0.8767	0.7690	**0.9182**	**0.9181**
		JP2K	0.8534	0.8883	0.8927	0.8097	**0.9009**	**0.9013**	0.8890	0.7895	0.8583	0.8988	0.8830	0.8680
		All	0.7820	0.7673	0.8401	0.8594	0.8787	0.8681	0.7972	0.8055	0.8417	0.7046	**0.8990**	**0.8937**
	CSIQ	JPEG	0.7996	0.8986	0.9049	0.9223	–	–	0.8830	0.9016	0.8996	**0.9502**	0.9181	**0.9295**
		GB	0.8716	0.8766	0.9026	0.8964	–	–	0.8925	0.8362	0.8578	**0.9060**	**0.9036**	0.9002
		WN	0.8663	0.7597	0.9250	0.9220	–	–	0.8090	0.8222	0.8500	0.6664	**0.9313**	**0.9358**
		JP2K	0.8308	0.8951	0.8665	0.9050	–	–	**0.9065**	0.8699	0.9061	**0.9075**	0.8628	0.8620
		All	0.8284	0.8511	0.8993	**0.9143**	–	–	0.8710	0.8416	0.8803	0.7712	0.8957	**0.8999**
	SIQAD	JPEG	0.0818	0.4299	0.2703	0.4171	0.1843	0.3990	0.4422	0.1451	0.2919	**0.7149**	**0.7377**	0.7022
		GB	0.0870	0.4404	0.6318	0.7704	0.6497	0.2324	0.5266	0.6238	0.4556	0.6663	**0.8350**	**0.8349**
		WN	**0.8865**	0.6361	0.8238	0.8357	0.6623	0.6768	0.8245	0.8416	0.8143	**0.8648**	0.8453	0.8452
		JP2K	0.0928	0.2300	0.0169	0.2852	0.5721	0.4371	0.2458	0.1937	0.3837	0.4612	**0.6203**	**0.5728**
		All	0.3575	0.2752	0.4164	0.6844	0.4934	0.3681	0.4816	0.5631	0.5167	0.4150	**0.7802**	**0.7714**
	Average		0.6438	0.6845	0.7267	0.7868	0.6900	0.6465	0.7857	0.7604	0.7865	0.8026	**0.8762**	**0.8710**
PLCC	LIVE	JPEG	–	–	–	–	–	–	0.9516	0.9437	0.9589	**0.9748**	0.9740	**0.9769**
		GB	–	–	–	–	–	–	**0.9446**	0.9112	**0.9327**	0.9150	0.9256	0.9237
		WN	–	–	–	–	–	–	0.9763	0.9280	0.9866	0.9645	**0.9874**	**0.9882**
		JP2K	–	–	–	–	–	–	**0.9264**	0.8658	0.9051	**0.9355**	0.9110	0.8934
		All	–	–	–	–	–	–	0.9162	0.8777	0.9164	0.8440	**0.9304**	**0.9320**
	TID2013	JPEG	0.6643	0.8774	0.8997	**0.9613**	0.9338	0.9181	0.8929	0.8693	0.8997	0.9536	0.9603	**0.9626**
		GB	0.8479	0.8492	0.8476	0.8494	**0.9214**	**0.8789**	0.8190	0.8478	0.8475	0.8355	0.8744	0.8627
		WN	0.8590	0.6480	0.8509	0.8759	0.7366	0.8189	0.8272	0.7972	0.8837	0.7749	**0.9262**	**0.9265**
		JP2K	0.9057	**0.9203**	0.9178	0.8587	**0.9263**	0.9347	0.9066	0.8093	0.8896	0.9163	0.9013	0.8890
		All	0.7859	0.7912	0.8662	0.8764	**0.8907**	**0.8901**	0.8091	0.8051	0.8576	0.8114	0.8895	0.8812
	CSIQ	JPEG	0.8239	0.9377	0.9463	0.9678	–	–	0.9347	0.9377	0.9546	**0.9693**	**0.9706**	0.9689
		GB	0.8993	0.8930	**0.9275**	0.9218	–	–	0.9249	0.8565	0.8937	**0.9298**	0.9182	0.9161
		WN	0.8878	0.7743	0.9376	0.9247	–	–	0.8113	0.8781	0.8638	0.6873	**0.9445**	**0.9476**
		JP2K	0.8962	0.9148	0.8972	**0.9379**	–	–	**0.9264**	0.8951	**0.9264**	0.9183	0.8916	0.8911
		All	0.8556	0.8792	**0.9240**	**0.9399**	–	–	0.8880	0.8736	0.9070	0.8657	0.9175	0.9188
	SIQAD	JPEG	0.1460	0.4696	0.2856	0.4221	0.2726	0.4017	0.4473	0.3687	0.4013	0.7302	**0.7525**	**0.7332**
		GB	0.4632	0.4585	0.6597	0.7579	0.6834	0.4296	0.6066	0.6255	0.5505	0.6551	**0.8462**	**0.8466**
		WN	0.8869	0.6415	0.8478	0.8537	0.6763	0.6843	0.8339	0.8526	0.8147	0.8757	**0.8927**	**0.8933**
		JP2K	0.1438	0.3166	0.2775	0.3176	**0.5975**	0.4708	0.3752	0.2448	0.4752	0.5359	**0.6297**	0.5804
		All	0.4043	0.3340	0.4962	0.7271	0.4983	0.3887	0.4996	0.5955	0.5400	0.4766	**0.7982**	**0.7930**
	Average		0.6980	0.7137	0.7721	0.8128	0.7137	0.6816	0.8109	0.7892	0.8202	0.8285	**0.8921**	**0.8863**
RMSE	LIVE	JPEG	–	–	–	–	–	–	9.7881	10.534	9.0327	**7.1003**	7.2118	**6.8035**
		GB	–	–	–	–	–	–	**6.0625**	7.6091	**6.6621**	7.4539	6.9904	7.0758
		WN	–	–	–	–	–	–	6.0507	10.423	4.5715	7.3908	**4.4300**	**4.2847**
		JP2K	–	–	–	–	–	–	**9.5003**	12.624	10.729	**8.9108**	10.403	11.337
		All	–	–	–	–	–	–	10.833	12.953	10.822	14.498	**9.9065**	**9.7964**
	TID2013	JPEG	1.1256	0.7226	0.6573	**0.4151**	0.5390	0.5968	0.6780	0.7445	0.6575	0.4536	0.4199	**0.4081**
		GB	0.6616	0.6589	0.6622	0.6585	**0.4850**	**0.5952**	0.7160	0.6618	0.6623	0.6858	0.6056	0.6310
		WN	0.3631	0.5401	0.3725	0.3421	0.4796	0.4070	0.3985	0.4281	0.3319	0.4482	**0.2674**	**0.2669**
		JP2K	0.7218	0.6661	0.6762	0.8728	**0.6417**	**0.6054**	0.7188	1.0004	0.7777	0.6819	0.7377	0.7797
		All	0.8626	0.8530	0.6971	0.6717	**0.6340**	**0.6357**	0.8197	0.8273	0.7174	0.8153	0.6373	0.6593
	CSIQ	JPEG	0.1734	0.1063	0.0989	0.0770	–	–	0.1088	0.1063	0.0912	**0.0752**	**0.0736**	0.0757
		GB	0.1253	0.1290	**0.1071**	0.1111	–	–	0.1090	0.1479	0.1286	**0.1055**	0.1135	0.1149
		WN	0.0772	0.1062	0.0583	0.0639	–	–	0.0981	0.0803	0.0845	0.1219	**0.0551**	**0.0536**
		JP2K	0.1402	0.1277	0.1396	**0.1096**	–	–	**0.1190**	0.1409	**0.1190**	0.1251	0.1431	0.1434
		All	0.1463	0.1346	**0.1080**	**0.0965**	–	–	0.1300	0.1375	0.1190	0.1415	0.1124	0.1116
	SIQAD	JPEG	9.2958	8.2959	9.0051	8.5182	9.0406	8.6050	8.4041	8.7345	8.6065	6.4203	**6.1880**	**6.3894**
		GB	13.450	13.487	11.405	9.9010	11.079	13.704	12.065	11.841	12.669	11.467	**8.0861**	**8.0781**
		WN	6.8920	11.443	7.9098	7.7688	10.988	10.877	8.2320	7.7941	8.6487	7.2015	**6.7208**	**6.7047**
		JP2K	10.285	9.8587	9.9852	9.8552	**8.3342**	9.1696	9.6338	10.077	9.1446	8.7752	**8.0742**	8.4633
		All	12.815	13.207	12.165	9.6200	12.148	12.910	12.138	11.257	11.794	12.318	**8.4399**	**8.5368**
#Hit			1/47	1/47	4/47	7/47	10/32	6/32	9/62	0/62	4/62	16/62	**34/62**	**34/62**
Time (seconds/image)			8.5404	22.951	0.7183	23.773	2.8918	0.2198	0.1072	**0.0426**	3.4352	**0.0144**	0.4959	0.5006

TABLE II
SRCC PERFORMANCE ON THE TID2013 DATABASE

Distortion Types	DIIVINE	BLIINDS-II	BRISQUE	NFERM	CORNIA	HOSA	NIQE	QAC	IL-NIQE	LPSI	BPRI(c)	BPRI(p)
Additive Gaussian noise	0.8553	0.6468	0.8520	0.8581	0.7354	0.8172	0.8187	0.7427	0.8767	0.7690	**0.9182**	**0.9181**
Additive noise in color components	0.7120	0.4762	0.7089	0.7096	0.7076	0.7534	0.6701	0.7184	0.8159	0.4955	**0.8600**	**0.8587**
Spatially correlated noise	0.4626	0.5862	0.4916	0.2184	0.6892	0.5812	0.6659	0.1694	**0.9233**	**0.6968**	0.4718	0.5293
Masked noise	0.6752	0.6183	0.5767	0.2210	0.7141	0.5565	**0.7464**	0.5927	0.5134	0.0462	0.7381	**0.7479**
High frequency noise	0.8778	0.7229	0.7526	0.8813	0.7972	0.8650	0.8454	0.8628	0.8691	0.9250	**0.9285**	**0.9263**
Impulse noise	**0.8063**	0.6525	0.6289	0.1728	0.7634	0.5592	0.7446	**0.8003**	0.7556	0.4324	0.4547	0.4585
Quantization noise	0.1650	0.7370	0.7932	0.7747	0.0922	0.6794	0.8514	0.7089	**0.8721**	**0.8537**	0.4890	0.4898
Gaussian blur	0.8344	0.8367	0.8137	0.8501	**0.9274**	0.8695	0.7986	0.8464	0.8148	0.8408	**0.8725**	0.8593
Image denoising	0.7231	0.6884	0.5849	0.6389	**0.8459**	**0.8453**	0.5900	0.3381	0.7494	0.2487	0.4322	0.4210
JPEG compression	0.6288	0.8360	0.8448	0.8722	0.8958	0.8957	0.8468	0.8369	0.8340	**0.9123**	0.9067	**0.9107**
JPEG2000 compression	0.8534	0.8883	0.8927	0.8097	**0.9009**	**0.9013**	0.8890	0.7895	0.8583	0.8988	0.8830	0.8680
JPEG transmission errors	0.2387	0.1098	0.3163	0.1322	0.6991	0.6552	0.0006	0.0491	0.2819	0.0911	**0.7359**	**0.7887**
JPEG2000 transmission errors	0.0606	**0.6409**	0.3595	0.1684	**0.6762**	0.3834	0.5114	0.4065	0.5240	0.6106	0.4431	0.4883
Non eccentricity pattern noise	0.0598	0.0997	0.1459	0.0646	**0.2332**	**0.1764**	0.0682	0.0477	0.0808	0.0520	0.0046	0.0086
Local block-wise distortions	0.0928	0.2440	0.2233	0.2020	0.2287	**0.2688**	0.1218	**0.2474**	0.1334	0.1372	0.2367	0.2333
Mean shift	0.0104	0.0963	0.1241	0.0213	0.0844	0.1276	0.1639	**0.3059**	0.1840	**0.3409**	0.0942	0.1106
Contrast change	**0.4601**	0.0011	0.0404	**0.2178**	0.1814	0.1377	0.0171	0.2067	0.0136	0.1992	0.1983	0.1846
Change of color saturation	0.0684	0.0119	0.1126	0.3067	0.0353	0.0479	0.2481	**0.3683**	0.1655	0.3018	0.2982	**0.3786**
Multiplicative Gaussian noise	0.7873	0.6193	0.7242	0.7162	0.6574	0.7314	0.6934	0.7902	0.6936	0.6959	**0.8620**	**0.8612**
Comfort noise	0.1156	0.1663	0.0076	0.1427	**0.5235**	**0.3658**	0.1544	0.1521	0.3614	0.0181	0.0973	0.0691
Lossy compression of noisy images	0.6327	0.4552	0.6856	0.6541	**0.8654**	0.7266	0.8023	0.6395	**0.8287**	0.2356	0.5975	0.5977
Color quantization with dither	0.4362	0.7677	0.7652	0.4790	0.3919	0.8017	0.7881	**0.8731**	0.7504	**0.8998**	0.6797	0.6753
Chromatic aberrations	0.6608	0.6445	0.6166	0.6430	**0.8183**	0.7209	0.5671	0.6249	0.6793	0.6953	0.7248	**0.7253**
Sparse sampling and reconstruction	0.8334	0.8257	0.7841	0.7847	0.8536	0.8564	0.8340	0.7856	**0.8643**	**0.8620**	0.7313	0.7873
Average	0.5021	0.5155	0.5352	0.4808	0.5966	**0.5968**	0.5599	0.5376	**0.6018**	0.5108	0.5691	0.5790
#Hit	2	1	0	1	**8**	6	1	5	5	6	6	**9**

methods perform pretty well on SCIs and show good consistency across image content types. It is not surprising since most current BIQA models are implicitly designed for NSIs and rely on some statistics of NSIs, while the quality features utilized by the proposed PRI-based methods are not restricted to natural scenes.

C. Performance on Non-Common Distortions

We test all BIQA models' generalizability to other distortions from 3 aspects: the whole TID2013 database which includes 24 distortion types; the CCT database which includes 3 content types and 2 distortion types; and non-common distortions in the LIVE, CSIQ, and SIQAD databases including fastfading (FF), additive pink Gaussian noise (pWN), contrast change (CC), motion blur (MB), and layer-segmentation based compression (LSC). The corresponding performance comparison results are listed in Tables II–IV respectively. We only report SRCC performance for simplicity. Similar results can be obtained according to other criteria.

From Table II, it is observed that CORNIA, HOSA, IL-NIQE and the proposed BPRI methods are the top models on the TID2013 according to the average performance and the hit count. The BPRI methods generalize to other distortion types well though they are integrated by three PRI-based distortion-specific metrics. On the CCT database, both the NSI, CGI, SCI subsets and the whole database are tested. It can be observed from Table III that the proposed BPRI methods are close to the best performing BIQA models on the NSI and CGI subsets, but they perform the best on the SCI subset and the whole database. The proposed BPRI methods demonstrate their superiority on SCIs, which is consistent with the results described in Section V-B. From an overall perspective, the BPRI methods still perform the best on the CCT database. From Table IV, we can observe that no BIQA model shows significant superiority on non-common distortions of the LIVE, CSIQ, and SIQAD databases. Most BIQA models show good performance on distortions close to the common distortions, but they are not good at distinctive distortions such as contrast change.

TABLE III
SRCC PERFORMANCE ON THE CCT DATABASE

Models	NSI	CGI	SCI	All	Average	#Hit
DIIVINE	0.5383	0.5995	0.1575	0.3634	0.4147	0
BLIINDS-II	0.7216	0.6662	0.2017	0.2424	0.4580	0
BRISQUE	0.5866	0.6182	0.5241	0.1247	0.4634	0
NFERM	0.6845	0.6745	0.3512	0.2382	0.4871	0
CORNIA	0.6843	0.7573	0.2174	0.4700	0.5323	0
HOSA	**0.7692**	0.7457	0.0239	0.2895	0.4571	1
NIQE	0.6693	0.6911	0.2923	0.2505	0.4758	0
QAC	**0.7511**	**0.8226**	0.0257	0.3783	0.4944	2
IL-NIQE	0.5674	0.6463	0.3453	0.1599	0.4297	0
LPSI	0.7487	**0.7705**	0.0385	0.4474	0.5013	1
BPRI(c)	0.7453	0.7626	**0.5751**	**0.4768**	**0.6400**	**3**
BPRI(p)	0.7392	0.7618	**0.6352**	**0.4815**	**0.6544**	**3**

From Tables II–IV, we can observe that BPRI works for various distortions though BPRI only considers blockiness, sharpness and noisiness. It is because that blockiness, blurring and noiseness are the most common and dominant distortions in both IQA databases and practical visual communication systems. Many other distortions are close to or can be combinations of these distortions. The mechanism of BPRI is

TABLE IV
SRCC PERFORMANCE ON NON-COMMON DISTORTIONS OF THE LIVE, CSIQ, AND SIQAD DATABASES

Models	LIVE	CSIQ		SIQAD			#Hit
	FF	pWN	CC	MB	CC	LSC	
DIIVINE	–	0.1766	0.3958	**0.4743**	0.1300	0.0206	1
BLIINDS-II	–	0.2011	0.0220	0.2512	0.0891	0.2077	0
BRISQUE	–	0.2516	0.0288	0.4401	0.0024	0.2470	0
NFERM	–	0.6262	0.3770	0.4238	**0.2826**	0.3008	1
CORNIA	–	–	–	0.2165	0.1893	0.2242	0
HOSA	–	–	–	0.2257	0.1530	0.2685	0
NIQE	**0.8630**	0.2973	0.2317	0.3514	0.0641	0.3483	1
QAC	0.8231	0.0019	0.2446	0.3755	0.0745	0.1866	0
IL-NIQE	0.8329	**0.8738**	0.4998	0.4480	0.0459	0.1567	1
LPSI	0.7808	0.2486	**0.5386**	0.3940	0.0676	0.5485	1
BPRI(c)	0.8207	0.3787	0.1076	0.0658	0.1720	**0.7479**	1
BPRI(p)	0.8181	0.3887	0.1563	0.0821	0.1656	0.7466	0

similar to DIIVINE, which also utilizes a 2-stage quality regression after distortion identification framework.

D. Parameter Sensitivity

The proposed PRI-based metrics (PSS, LSS_s, LSS_n and BPRI) involve very few parameters. Most parameters are introduced when generating the PRI, where we need to control the distortion level of the PRI. The key parameters include:

- The compression "$\mathrm{Quality}$" parameter in PSS, which specifies the compression degree when deriving the PRI.
- The blurring filter in LSS_s, which specifies the blurring degree when blurring the distorted image to the PRI.
- The variance v in LSS_n, which specifies the intensity of the added noise when deriving the PRI.
- The choice of FR metric in BPRI, which is used to guide the score alignment.

We test the BPRI's sensitivity to variations of these parameters on the 4 common distortions (JPEG, GB, WN and JP2K) of the LIVE, TID2013, CSIQ and SIQAD databases. When testing one parameter, other parameters are fixed as the default settings. Since the classifier and the score alignment are related to all quality scores of the PRI-based distortion-specific metrics, we re-train the classifier and re-fit the score alignment model when testing the parameters. For the second parameter, we replace the averaging filter with a 3×3 Gaussian filter with standard deviation (STD) of σ, and test BPRI's sensitivity to σ. The averaging filter can be approximated by a Gaussian filter with a large σ.

Fig. 7 illustrates the SRCC performance of BPRI(p) with the variations of the first three parameters. It can be observed that the performance remains stable within a significantly wide range, which suggests high generalization capability of the proposed BPRI method. Note that these three parameters have controlled the distortion levels of the PRIs. It means that the PRI does not need to be located at an exact point in the quality axis as illustrated in Fig. 1. The PRI can be located at any point within a certain range which represents worse quality than the distorted image. Table V lists the SRCC performance of the BPRI(p) using different FR metrics to guide the score alignment. We have tested several representative FR metrics including: SSIM [53], MS-SSIM [54], VIF [55], FSIM [56], VSI [57], MDSI [58], PSIM [59] and GMSD [49]. It is observed that the performance of using different FR metrics are quite close, especially when using VIF, FSIM, PSIM and GMSD.

TABLE V
SRCC PERFORMANCE OF THE BPRI USING DIFFERENT FR METRICS TO GUIDE THE SCORE ALIGNMENT

FR Metrics	LIVE	TID2013	CSIQ	SIQAD	Average
SSIM	0.8900	0.8145	0.8832	0.6998	0.8219
MS-SSIM	0.9120	0.8695	0.8972	0.7229	0.8504
VIF	0.9294	0.8855	0.8884	0.7784	0.8704
FSIM	0.9158	0.8789	0.9025	0.7543	0.8629
VSI	0.9076	0.8407	0.8934	0.6720	0.8284
MDSI	0.9210	0.8542	0.8885	0.6605	0.8311
PSIM	0.9295	0.8860	0.8964	0.7378	0.8624
GMSD	0.9304	0.8937	0.8999	0.7714	0.8739

E. Computational Complexity

To compare the computational complexity of state-of-the-art BIQA models, we report the average running time (seconds/image) for 100 images with a fixed resolution of 512×512 on a computer with 4.00 GHz Intel Core i7-6700K CPU and 32 GB RAM. The running cost includes all feature extraction and quality prediction time. From Table I, it appears that BPRI is among the models with low computational complexity.

F. Analyses and Discussions

As described above, the proposed PRI-based BPRI method achieves competitive performance compared with the state-of-the-art. It is interesting to analyse how the PRI contributes to the BIQA process. Fig. 8 illustrates example images of different distortion types and levels, and the corresponding PRIs. Three distortion levels and three distortion types (JPEG, WN and GB) are illustrated here.

Taking JPEG compression as an example, we can observe that no matter what the distorted images look like, the PRIs look very similar, especially in the feature space (for example, the pseudo structure space for JPEG images). That is the reason why the "pseudo-reference" can be used as a "reference". The distorted image becomes more and more similar to the PRI as the distortion becomes severer and severer. Different from the traditional "reference" which describes image contents with perfect quality, PRI describes the image contents under the extremely poor quality condition, and the PRI is independent of original distortion level to some extent. One bottleneck of traditional BIQA models is that it is not easy for them to distinguish artifacts and image contents, which are sometimes similar. PRI-based methods have considered the influence of image contents by comparing with the PRI.

Similarly, PRI plays the same role in the QA of noisy images. PRI provides a description of the image content under extremely noisy condition. The PRI is image content dependent, but it is

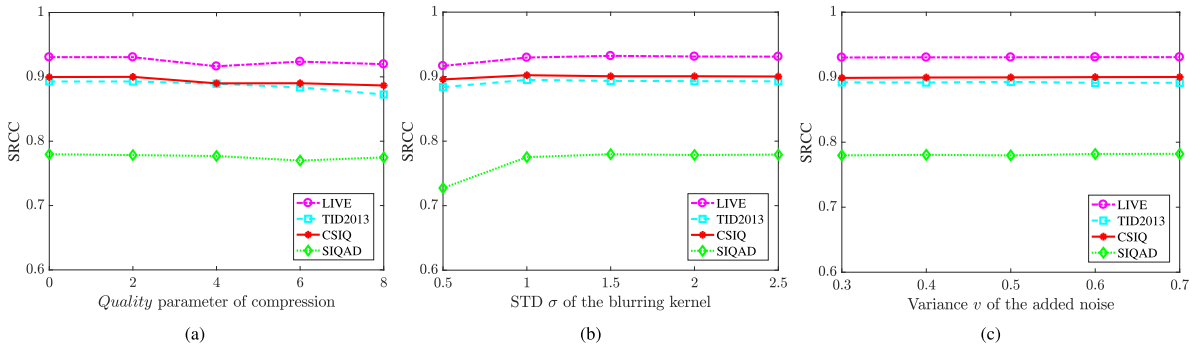


Fig. 7. The SRCC performance of BPRI w.r.t. variations of (a) the compression quality parameter, (b) the STD of the blurring kernel, (c) the variance of the added noise on the LIVE, TID2013, CSIQ and SIQAD databases.

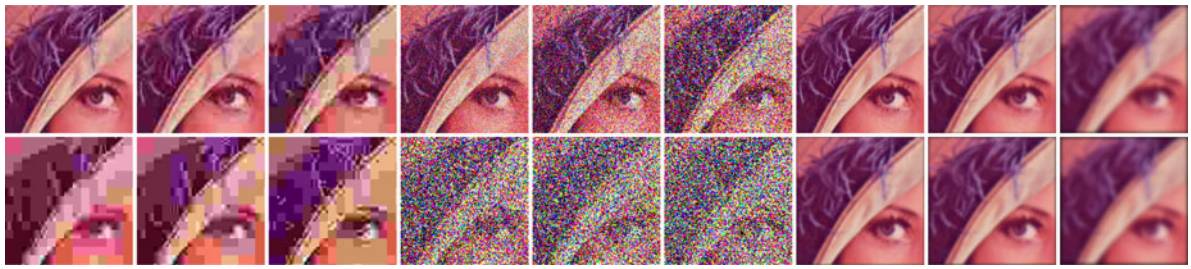


Fig. 8. Example images of different distortion types (from left to right: JPEG, WN, GB) and levels (three levels for each type), and the corresponding PRIs. Top: distorted images; down: the corresponding PRIs.

independent of original noise level to some extent. The philosophy of PRI-based sharpness estimation is slightly different from PRI-based blockiness and noisiness estimation. We introduce certain amount of blurring (e.g., 3×3 averaging in our work), and estimate the image content change. Blurring can change image content more significantly in sharper images. The PRI acts as a reference to quantify the content change.

The most critical engineering design is the deriving of the PRI. Since each type of distortion causes specific artifacts, we need to introduce the same type of artifacts to be consistent when deriving the PRI. Thus the way of introducing artifacts is determined, i.e., JPEG compression, adding noise, or blurring. We only need to control the strength of the introduced artifacts when deriving the PRI. We have introduced different levels of artifacts in Section V-D. As analysed, the performance is stable within a wide range, which indicates that the PRI does not need to be located at an exact point in the quality axis as illustrated in Fig. 1. The quality conditions of the PRI or the strength of the introduced artifacts are adjustable, and they will not influence the quality prediction significantly within a certain range.

VI. Conclusion

Contrary to traditional IQA metrics which explicitly or implicitly measure the distorted image's deviation from perfect quality images, we have utilized in this paper a new concept of pseudo-reference image (PRI), and a PRI-based BIQA framework by measuring the distorted image's distance to the PRI. We confirm the effectiveness of PRI-based IQA, and propose several PRI-based distortion-specific metrics to estimate blockiness, sharpness and noisiness. The PRI-based distortion-specific metrics are then integrated into a general-purpose BIQA method named blind PRI-based (BPRI) metric. The proposed BPRI method is opinion-unaware and almost training-free except for the distortion identification process. Compared with the state-of-the-art competitors, the proposed model shows superior or at least comparable quality prediction performance.

One advantage of the proposed metrics is that they perform quite well on both natural scene and screen content images. It is an important and useful characteristic since the computer-generated screen contents become more and more widespread. Note that although we only predict image quality via estimating blockiness, sharpness and noisiness in this research, the presented PRI and PRI-based IQA framework could be generalized to other image distortion types, and the proposed BPRI method could also be further improved by integrating more PRI-based distortion-specific estimators. For other distortion types, we need to work out specific problems of PRI definition and appropriate distance measures. We are currently working on more extensions of the introduced PRI scheme to the assessment of more distortion types.

References

[1] X. Min *et al.*, "Blind quality assessment of compressed images via pseudo structural similarity," in *Proc. IEEE Int. Conf. Multimedia Expo*, 2016, pp. 1–6.

[2] W. Lin and C.-C. J. Kuo, "Perceptual visual quality metrics: A survey," *J. Vis. Commun. Image Represent.*, vol. 22, no. 4, pp. 297–312, 2011.

[3] Z. Wang and A. C. Bovik, "Mean squared error: Love it or leave it? A new look at signal fidelity measures," *IEEE Signal Process. Mag.*, vol. 26, no. 1, pp. 98–117, Jan. 2009.

[4] Z. Wang and A. C. Bovik, "Reduced-and no-reference image quality assessment," *IEEE Signal Process. Mag.*, vol. 28, no. 6, pp. 29–40, Nov. 2011.

[5] F. Crete, T. Dolmiere, P. Ladret, and M. Nicolas, "The blur effect: Perception and estimation with a new no-reference perceptual blur metric," *Proc. SPIE*, vol. 6492, 2007, Art. no. 64920I.

[6] 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, 2011.

[7] 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.

[8] 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.

[9] H. R. Sheikh, Z. Wang, L. Cormack, and A. C. Bovik, "LIVE image quality assessment database release 2." [Online]. Available: http://live.ece.utexas.edu/research/quality.

[10] N. Ponomarenko *et al.*, "Image database TID2013: Peculiarities, results and perspectives," *Signal Process., Image Commun.*, vol. 30, pp. 57–77, 2015.

[11] 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.

[12] 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.

[13] 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.

[14] A. C. Bovik and S. Liu, "DCT-domain blind measurement of blocking artifacts in DCT-coded images," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, vol. 3, 2001, pp. 1725–1728.

[15] Z. Wang, H. R. Sheikh, and A. C. Bovik, "No-reference perceptual quality assessment of JPEG compressed images," in *Proc. IEEE Int. Conf. Image Process.*, vol. 1, 2002, pp. I-477–I-480.

[16] H. Liu and I. Heynderickx, "A perceptually relevant no-reference blockiness metric based on local image characteristics," *EURASIP J. Adv. Signal Process.*, vol. 2009, 2009, Art. no. 263540.

[17] S. A. Golestaneh and D. M. Chandler, "No-reference quality assessment of JPEG images via a quality relevance map," *IEEE Signal Process. Lett.*, vol. 21, no. 2, pp. 155–158, Feb. 2014.

[18] L. Li, W. Lin, and H. Zhu, "Learning structural regularity for evaluating blocking artifacts in JPEG images," *IEEE Signal Process. Lett.*, vol. 21, no. 8, pp. 918–922, Aug. 2014.

[19] Y. Zhan and R. Zhang, "No-reference JPEG image quality assessment based on blockiness and luminance change," *IEEE Signal Process. Lett.*, vol. 24, no. 6, pp. 760–764, Jun. 2017.

[20] R. Ferzli and L. J. Karam, "A no-reference objective image sharpness metric based on the notion of just noticeable blur (JNB)," *IEEE Trans. Image Process.*, vol. 18, no. 4, pp. 717–728, Apr. 2009.

[21] N. D. Narvekar and L. J. Karam, "A no-reference image blur metric based on the cumulative probability of blur detection (CPBD)," *IEEE Trans. Image Process.*, vol. 20, no. 9, pp. 2678–2683, Sep. 2011.

[22] L. Li *et al.*, "Image sharpness assessment by sparse representation," *IEEE Trans. Multimedia*, vol. 18, no. 6, pp. 1085–1097, Jun. 2016.

[23] P. V. Vu and D. M. Chandler, "A fast wavelet-based algorithm for global and local image sharpness estimation," *IEEE Signal Process. Lett.*, vol. 19, no. 7, pp. 423–426, Jul. 2012.

[24] C. T. Vu, T. D. Phan, and D. M. Chandler, "S_3: A spectral and spatial measure of local perceived sharpness in natural images," *IEEE Trans. Image Process.*, vol. 21, no. 3, pp. 934–945, Mar. 2012.

[25] R. Hassen, Z. Wang, and M. M. Salama, "Image sharpness assessment based on local phase coherence," *IEEE Trans. Image Process.*, vol. 22, no. 7, pp. 2798–2810, Jul. 2013.

[26] K. Gu, G. Zhai, W. Lin, X. Yang, and W. Zhang, "No-reference image sharpness assessment in autoregressive parameter space," *IEEE Trans. Image Process.*, vol. 24, no. 10, pp. 3218–3231, Oct. 2015.

[27] D. Zoran and Y. Weiss, "Scale invariance and noise in natural images," in *Proc. IEEE 12th Int. Conf. Comput. Vis.*, 2009, pp. 2209–2216.

[28] C. Liu, W. T. Freeman, R. Szeliski, and S. B. Kang, "Noise estimation from a single image," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, vol. 1, 2006, pp. 901–908.

[29] C. Tang, X. Yang, and G. Zhai, "Noise estimation of natural images via statistical analysis and noise injection," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 25, no. 8, pp. 1283–1294, Aug. 2015.

[30] G. Zhai, A. Kaup, J. Wang, and X. Yang, "A dual-model approach to blind quality assessment of noisy images," *APSIPA Trans. Signal Inf. Process.*, vol. 4, 2015, Art. no. e4.

[31] M. Liu *et al.*, "Blind image quality assessment for noise," in *Proc. IEEE Broadband Multimedia Syst. Broadcast.*, 2014, pp. 1–5.

[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] 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.

[34] 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.

[35] 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.

[36] 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.*, 2012, pp. 1098–1105.

[37] 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.

[38] 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.

[39] 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.

[40] 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.

[41] W. Xue, L. Zhang, and X. Mou, "Learning without human scores for blind image quality assessment," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2013, pp. 995–1002.

[42] Q. Wu, Z. Wang, and H. Li, "A highly efficient method for blind image quality assessment," in *Proc. IEEE Int. Conf. Image Process.*, 2015, pp. 339–343.

[43] J. Shi and C. Tomasi, "Good features to track," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 1994, pp. 593–600.

[44] J. Wu *et al.*, "Visual orientation selectivity based structure description," *IEEE Trans. Image Process.*, vol. 24, no. 11, pp. 4602–4613, Nov. 2015.

[45] X. Min, G. Zhai, Z. Gao, and K. Gu, "Visual attention data for image quality assessment databases," in *Proc. IEEE Int. Symp. Circuits Syst.*, 2014, pp. 894–897.

[46] 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. 6:1–6:23, 2016.

[47] X. Min *et al.*, "Visual attention analysis and prediction on human faces," *Inf. Sci.*, vol. 420, pp. 417–430, 2017.

[48] K. Ma *et al.*, "Waterloo exploration database: New challenges for image quality assessment models," *IEEE Trans. Image Process.*, vol. 26, no. 2, pp. 1004–1016, Feb. 2017.

[49] 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.

[50] C.-C. Chang and C.-J. Lin, "LIBSVM: A library for support vector machines," *ACM Trans. Intell. Syst. Technol.*, vol. 2, no. 3, 2011, Art. no. 27.

[51] 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, 2018.

[52] K. Gu *et al.*, "Evaluating quality of screen content images via structural variation analysis," *IEEE Trans. Vis. Comput. Graph.*, to be published.

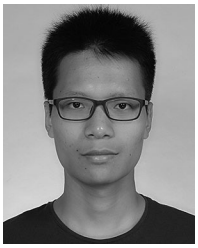
[53] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.

[54] 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, 2003, pp. 1398–1402.

[55] 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.

[56] 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.

[57] L. Zhang, Y. Shen, and H. Li, "VSI: A visual saliency-induced index for perceptual image quality assessment," *IEEE Trans. Image Process.*, vol. 23, no. 10, pp. 4270–4281, Oct. 2014.
[58] H. Z. Nafchi, A. Shahkolaei, R. Hedjam, and M. Cheriet, "Mean deviation similarity index: Efficient and reliable full-reference image quality evaluator," *IEEE Access*, vol. 4, pp. 5579–5590, 2016.
[59] 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.



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

Ke Gu received the B.S. and Ph.D. degrees in electronic engineering from Shanghai Jiao Tong University, Shanghai, China, in 2009 and 2015. His research interests include quality assessment, contrast enhancement, visual saliency detection, and air quality prediction. Dr. Gu is an Associate Editor for the IEEE ACCESS, and is 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, the *Journal of Selected Topics in Signal Processing*, the IEEE SIGNAL PROCESSING LETTERS, IEEE ACCESS, *Information Sciences*, *Neurocomputing*, *Signal Processing: Image Communication*, the *Journal of Visual Communication and Image Representation*, *Digital Signal Processing*, *Multimedia Tools and Applications*, ELL, etc. He has reviewed more than 50 journal papers each year. He was the 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 the leading Special Session Organizer in VCIP2016 and ICIP2017.

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. He is currently a Research Professor with the Institute of Image Communication and Information Processing, Shanghai Jiao Tong University. From 2008 to 2009, he was a Visiting Student at 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, Erlangen, Germany. His research interests include multimedia signal processing and perceptual signal processing. Prof. Zhai was the recipient of the Award of National Excellent Ph.D. thesis from the Ministry of Education of China in 2012.

Jing Liu received the B.E. and Ph.D. degrees from Shanghai Jiao Tong University, Shanghai, China, in 2011 and 2017, respectively. She is currently an Assistant Professor with the Multimedia Institute, Tianjin University, Tianjin, China. From 2014 to 2015, she was a Visiting Student at the Department of Computer Science and Engineering, State University of New York at Buffalo, NY, USA. Her research interests include multimedia signal processing and perceptual visual processing.

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, China, 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 visited the Institute for Computer Science, University of Freiburg, Freiburg im Breisgau, Germany, as an Alexander von Humboldt Research Fellow. He has authored or co-authored more than 200 refereed papers, and has filed 60 patents. His current research interests include image processing and communication, computer vision, and machine learning. Dr. Yang is an Associate Editor for the IEEE TRANSACTIONS ON MULTIMEDIA and a Senior Associate Editor for the IEEE SIGNAL PROCESSING LETTERS. He was a Series Editor of Springer *Communications in Computer and Information Science*, and an Editorial Board Member of *Digital Signal Processing*. He is a Member of 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. He is also the Chair of the Multimedia Big Data Interest Group of MMTC Technical Committee, IEEE Communication Society.

Chang Wen Chen (F'04) received the B.S. degree from the University of Science and Technology of China, Hefei, China, in 1983, the M.S.E.E. degree from the University of Southern California, Los Angeles, CA, USA, in 1986, and the Ph.D. degree from the University of Illinois at Urbana-Champaign, Champaign, IL, USA, in 1992. He is currently an Empire Innovation Professor with the Computer Science and Engineering, University at Buffalo, State University of New York, Buffalo, NY, USA. He was an Allen Henry Endow Chair Professor with the Florida Institute of Technology from 2003 to 2007. He was on the faculty of the Electrical and Computer Engineering, University of Rochester from 1992 to 1996, and on the faculty of the Electrical and Computer Engineering, University of Missouri-Columbia, from 1996 to 2003. Dr. Chen has been the Editor-in-Chief for the IEEE TRANSACTIONS ON MULTIMEDIA since 2014. He has also served as the Editor-in-Chief for the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY from 2006 to 2009. He has been an Editor for several other major IEEE transactions and journals, including the PROCEEDINGS OF THE IEEE, the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS, and the IEEE JOURNAL ON EMERGING AND SELECTED TOPICS IN CIRCUITS AND SYSTEMS. He has served as the Conference Chair for several major IEEE, ACM, and SPIE conferences related to multimedia, video communications, and signal processing. His research is supported by NSF, DARPA, Air Force, NASA, Whitaker Foundation, Microsoft, Intel, Kodak, Huawei, and Technicolor. He was the recipient of several research and professional achievement awards, including the Sigma Xi Excellence in Graduate Research Mentoring Award in 2003, the Alexander von Humboldt Research Award in 2010, the University at Buffalo Exceptional Scholar-Sustained Achievement Award in 2012, and the State University of New York Chancellor's Award for Excellence in Scholarship and Creative Activities in 2016. He and his students have received eight best paper awards or best student paper awards over the past two decades. He is an SPIE Fellow.