Signal Processing: Image Communication 79 (2019) 54–62

Contents lists available at ScienceDirect

Signal Processing: Image Communication

journal homepage: www.elsevier.com/locate/image

# Quality assessment of images with multiple distortions based on phase congruency and gradient magnitude☆

Xikui Miao [a,b,*], Hairong Chu [c], Hui Liu [c], Yao Yang [d], Xiaolong Li [a]

[a] School of Information Engineering, Henan University of Science and Technology, Luoyang, Henan, 471000, China
[b] Department of Electrical and Computer Engineering, Brigham Young University, Provo, UT, 84602, USA
[c] Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences, Changchun, Jilin, 130033, China
[d] School of Astronautics, Northwestern Polytechnical University, Xi'an, Shaanxi, 710072, China

ARTICLE INFO

*Keywords:*
No-reference
Image quality assessment
Phase congruency
Local binary pattern
Image gradient

ABSTRACT

In image communication systems, images are often contaminated by multiple types of distortions. However, most existing image quality assessment (IQA) methods mainly focused on a single type of distortions. In this paper, we proposed a no-reference (NR) IQA method for images with multiple distortions. Image distortions not only destroy the intensity of low-level image features, but also alter their distribution, to both of which the human vision system (HVS) is sensitive. Based on these observations, low-level features are represented by phase congruency (PC) which is consistent with human perception. The distribution of low-level features is extracted using local binary pattern (LBP) in PC domain at multiple scales, which can effectively characterize the impact of multiple distortions on images. Given that PC is contrast invariant while the contrast does affect perceptual image quality of the HVS, image gradient magnitude (GM) is employed as a weighting factor for LBP histogram creation. Finally a support vector regression model is trained to map the gradient-weighted LBP histograms in PC domain at multi-scale to quality scores. Experimental results on two benchmark databases demonstrate that the proposed method achieves high consistency with subjective perception and performs better than other state-of-the-art full-reference (FF) and NR IQA methods.

## 1. Introduction

With the exponential growth of image and video data, there is a need to assess the image quality for any system that processes images and videos for human viewing [1]. The performance of image quality assessment (IQA) has become a critical metric for these systems such as image restoration, image/video compression [2], image forensics, and so on. During image acquisition, processing, transmission and storage, multiple types of distortions are introduced into the images (i.e., images are often distorted by more than two types of noises, such as Gaussian blur + additive white Gaussian + JPEG compression) [3,4]. Assessing the quality of such contaminated images effectively and efficiently has gained attention over the past several years.

Subjective quality assessment is the most reliable way to assess image quality, but it is labor-intensive and time-consuming. Therefore automatic IQA methods are in great demand. According to the availability of reference data, the objective IQA methods can be classified into three categories [5]: Full-Reference (FR), No-Reference (NR), and Reduced-Reference (RR). FR methods [6] require the whole reference image (high-quality) for quality assessment. RR methods are designed to predict the perceptual quality by using partial information of the reference image. However, the reference image or reference data is not always available in practice and NR methods [7,8] are expected to predict quality of distorted image without using any external information, making NR IQA an extremely challenging task. In this work, we mainly focus on developing an effective NR IQA.

Although many NR IQA methods are proposed, most of them are specially designed for images that are corrupted by only one of many possible distortions. However, in image communication systems, images usually undergo acquisition, processing, compression, transmission, storage, etc. In this pipeline, multiple types of distortions may be induced, resulting in contaminated images with blockiness and blur artifacts of JPEG compression, blurring due to processing and noise injection due to AD conversion [9]. Therefore, it is more practical to assess the quality for multiply-distorted images.

In fact, some of the state-of-the-art NR IQA methods show satisfactory performance on single distortion databases, such as TID2013 [10], LIVE [11] and CSIQ [12], in which the images are contaminated by a

☆ No author associated with this paper has disclosed any potential or pertinent conflicts which may be perceived to have impending conflict with this work. For full disclosure statements refer to https://doi.org/10.1016/j.image.2019.08.013.

* Corresponding author at: School of Information Engineering, Henan University of Science and Technology, Luoyang, Henan, 471000, China.
*E-mail address:* miaoxikui@163.com (X. Miao).

https://doi.org/10.1016/j.image.2019.08.013
Received 24 September 2018; Received in revised form 26 August 2019; Accepted 28 August 2019
Available online 30 August 2019

single type of distortions. But when these methods are confronted with the multiply-distorted databases, such as MLIVE [13], MDID2013 [14], they show very poor performance. Moreover, Chandler [15] pointed out that multiply-distorted images are a big challenge for IQA, because an IQA method must consider both the joint effects of distortions on the image and the effects of distortions on each other. In this paper, we focus on building an effective NR IQA method for multiple distortions.

It is well known that the human vision system (HVS) is highly adapted to extract structure for image perception and understanding [16] and is extremely sensitive to the changes of the structure [17]. Image structure is composed of low-level features, such as edges, lines, corners, zero-crossings and other local features, which consist of two aspects: intensity of structure and distribution of structure. Moreover, the distribution of these low-level features (structure) makes the visual content have its own distinctiveness. Visual content that low-level features contain is hidden behind the intensity of structure and distribution of structure. Accordingly, when an image is deteriorated, we believe that the intensity and distribution of low-level features would vary as a function of distortions (i.e., distortions would not only destroy the intensity of these low-level features but also alter the distribution of these low-level features). This motivates us to explore the use of low-level features and their distribution, on which the understanding of an image the HVS mainly depends, for image quality prediction.

However, what kinds of features could be used in designing an IQA method is an open question. According to phase congruency [18], low-level image features such as edges, lines, corners, zero-crossings and other local features all give rise to points where the Fourier components of the image are maximally in phase, and PC is a dimensionless measure of low-level feature significance. In other words, we can extract highly informative features of an image by PC approach. So in this work, PC map is employed to represent low-level image features. Given that PC is contrast invariant, but local image contrast does affect perceptual image quality of the HVS, image gradient magnitude (GM) is employed as complementary features for PC to represent the contrast information. PC and GM can reflect different aspects of the HVS in evaluating the quality of image, so they play complementary roles in characterizing the image features. Specifically, low-level image features are represented by PC, and their distribution is extracted by local binary pattern (LBP) histogram in PC domain at multiple scales. Then GM is employed as complementary feature for PC to build gradient-weighted LBP histogram, where the contrast degradation is considered. Finally a support vector regression model is trained to map histograms to quality scores, outperforming the state-of-the-art methods.

To the best of our knowledge, this is the first work that combining gradient magnitude (GM) and LBP in phase congruency domain at multiple scales to design the NR IQA method for multiply distorted images. Although PC and LBP are not new to IQA, the related methods often employ PC or LBP separately for full reference or reduced reference IQA, and only focus on a single type of possible distortions, making little contribution to multiply-distorted IQA which is regarded as a big challenge in NR IQA area. In this paper, we proposed that image distortions not only destroy the intensity of low-level features, but also alter their distribution, while other methods only notice one of the two aspects. Taking both into consideration, we developed an effective NR IQA method for multiply-distorted images, and the proposed method outperformed the state-of-the-art methods.

The rest of this paper is organized as follows. In Section 2, Related Work is presented. In Section 3, phase congruency and its computation are introduced, and then image gradient magnitude (GM) is extracted. In Section 4, our NR IQA method based on PC and GM is detailed. Experimental results and analysis of the proposed NR IQA method are presented in Section 5. Finally, conclusions are drawn in Section 6.

## 2. Related work

### 2.1. NR IQA methods for single distortion

The last several years have seen a surge of NR IQA methods for single distortion, which can be classified roughly into three categories. The first category of NR IQA methods for single distortion is mainly motivated by the characteristics of the HVS or the recent findings in brain science, such as free energy theory and orientation selectivity mechanism (OSM). Inspired by OSM, Wu [19] designed a new set of visual patterns for image content representation and proposed a new NR IQA based on the visual patterns. The image quality is predicted with a regression procedure. However, the computational complexity of this method is very high due to large number of the visual patterns. Based on the fact that the HVS is sensitive to luminance change and texture information for image perception, Fang [20] proposed a novel NR IQA method by incorporating statistical luminance and texture features for screen content images (SCIs) with both local and global feature representation. Zhai [21] developed RR free-energy-based distortion metric (FEDM) and NR free-energy-based quality metric (NFEQM). The premise of the two methods is that visual cognition is an active inference process, fitting brain model to visual sensory data. The quality of the image is quantified using the free energy. Gu [22] proposed NR Free energy and Structural degradation model based Distortion Metric (NFSDM) by integrating a pair of RR IQA algorithms (FEDM) [21] and structural degradation model (SDM) [23].

Inspired by NFSDM, Gu [24] extracted 23 features from the image, including free-energy-based features and the HVS-inspired features. A support vector regression (SVR) model is utilized to predict the image quality score (NFERM). Xie [25] introduced a novel application-driven IQA model for multiply-distorted dermoscopy images. Blur and uneven illumination are separately evaluated by taking the two single distortion levels as inputs. The overall image quality is predicted by a fuzzy neural network. Xue [26] utilized the joint statistics of the gradient magnitude (GM) and the Laplacian of Gaussian (LOG) responses. The image quality score is predicted by the SVR model trained by GM and LOG features (GMLOG). Ye [27] employed raw-image-patches as local descriptors and soft-assignment coding with max pooling to obtain effective image representation for quality estimation (CORNIA).

The second category of NR IQA methods for single distortion is targeted to predict image quality via natural scene statistics (NSS). Moorthy [28] proposed a new two-step framework for NR IQA based on NSS (BIQI), which identified the type of distortions in the image and then quantified the distortion using a trained SVR model. Based on two-step framework NR IQA, Moorthy [29] proposed NR IQA method called DIIVINE, relying on the hypothesis that natural scenes possess certain statistical properties which are altered in the presence of distortion. However, DIIVINE operates based on real-valued wavelet coefficients, whereas the visual appearance of an image can be strongly determined by both the magnitude and phase information. Saad [30] develop an efficient NR IQA method (BLIINDS-II) using NSS model of discrete cosine transform (DCT) coefficients based on Bayesian inference model. This method employed a probabilistic prediction model to predict image quality score with a small number of computationally convenient DCT-domain features.

Mittal [31] proposed NSS based distortion NR IQA method (BRISQUE) in spatial domain. The method used scene statistics of locally normalized luminance coefficients to quantify possible losses of ‘naturalness’ in the image due to the presence of distortions. Mittal [32] proposed NR IQA (NIQE) based on the construction of a ‘quality aware’ collection of NSS features extracted from the image domain. The image quality is expressed as the distance between a multivariate Gaussian (MVG) fit of the NSS features and a MVG model. Inspired by NIQE, Zhang [33] presented a complex extension of the DIIVINE (C-DIIVINE), which analyzed distorted images by using a complex version of the steerable pyramid wavelet transform. Three types of quality-aware

statistical features are extracted by the statistical features. C-DIIVINE blindly assessed image quality based on a complex Gaussian scale mixture. Zhang [34] extracted five types of NSS features from a collection of pristine naturalistic images and used them to learn a multivariate Gaussian (MVG) model of pristine images to predict the image quality (ILNIQE).

The third category of NR IQA methods for single distortion is devoted to specific distortion types. Li [35] proposed NR IQA method for deblocked images by simultaneously evaluating blocking artifacts and blur with a sharpness module and a blockiness module. This method evaluated the quality of deblocked images by combining the two modules. Considering that the current distortion-based IQA methods are very limited in the quality evaluation of enhanced images, Li [36] proposed a new quality method for enhanced images by simultaneously measuring non-structural information, sharpness and naturalness. A total of 42 perceptual features are extracted and used to train SVR model to predict the image quality. More recently, Li [37] presented NR IQA method for deblured images based on NSS features with SVR procedure. NSS features are extracted in both the spatial and frequency domains to account for both the global and local aspects of distortions in deblurred images.

Fang [38] proposed NR IQA method for contrast-distorted images based on NSS. This method extracted moment and entropy features from images and built NSS models upon them using a large-scale image database. The SVR model is adopted to map the features to image quality. Gu [39] devised a new NR IQA method of contrast distortion based on the concept of information maximization. The method generated an overall quality score of a contrast-distorted image by combining the entropy of unpredicted areas and uniformly distributed histogram. Li [40] proposed a novel NR image blur assessment metric, which is based on the observation that blur distortion changes the shape of an image. Shape changes can be represented using discrete Tchebichef moments, and image blur score is generated by normalizing the moment energy using the block variances with the guidance of a visual saliency model.

### 2.2. NR IQA methods for multiple distortions

Compared with IQA for single type distorted images, IQA for multiply-distorted images has thus far received less attention. Until now, research on NR IQA for multiply-distorted images has been very limited due to not only being a more challenging task, but also the lack of benchmark databases. And there are only a few papers that focus on NR IQA for multiply-distorted images. Inspired by the early human visual model and free energy based brain theory, Gu [14,41] proposed a SIx-Step BLInd Metric (SISBLIM) and a FIve-Step BLInd Metric(FISBLIM) for both singly and multiply distorted images and built a new multiply distorted image database (MDID2013) for IQA.

Li [42] proposed NR IQA method (GWH-GLBP) for multiply-distorted images based on structural degradation, using local binary pattern (LBP) to characterize image structure and predict the quality by measuring the degradation on image structure. The LBP feature was extracted from the image gradient map, which is effective to describe the complex degradation pattern introduced by multiple distortions. Lu [43] proposed NR IQA model for multiply-distorted images based on an improved Bag-of-Words (BoW) model using selected features. The features are selected from NSS features selected based on the correlation analysis, then an improved BoW model is applied to encode the selected features. Finally, the linear combination is used to map the features to the quality score.

Dai [44] proposed NR IQA method (BQASD, denoted by this paper) based on the fact that the HVS is sensitive to image structural information. Quality-aware features are extracted from both the first- and high-order image structures by LBP. The first-order feature set is extracted from the gradient maps of distorted images. The high-order feature set is calculated in the normalized luminance maps of distorted images. The SVR model is used for mapping two feature sets to quality score. Based on BQASD, Dai further improve the performance of IQA by making use of redundancy of features (first- and high-order feature sets) with random subspace method. With random subspace method, the proposed method [45] (IBQASD, denoted by this paper) can avoid overfitting. Li [46] proposed NR IQA (SHANIA) based on the statistical characterization in the shearlet domain and employed the most natural parts of an image to predict the natural tendency of other vulnerable parts. The predicted parts act as 'reference' and the difference between the reference and distorted parts is used as an indicator to predict the image quality.

## 3. Phase congruency and image gradient

According to the characteristics of the HVS, human eyes are highly adapted to extract low-level features for image perception [17]. To some extent, these low-level features can characterize image structures, which include intensity and distribution. It is the low-level features and their distribution that convey the main visual information. In this section, we introduce phase congruency which can be used to extract low-level image features. Considering that PC is contrast invariant, we propose using image gradient magnitude as complementary features to jointly represent image structures.

### 3.1. Phase congruency (PC) and its computation

Phase is an important signal component, which is often ignored in favor of magnitude, but phase is sufficient for image feature detection [18]. The significance of the phase information proved by the experiment of Oppenheim [47] played an important motivation factor for using PC as a low-level feature descriptor to represent image structures for IQA. This experiment showed that the phase of an image carries more structural information than the amplitude does. The PC approach can be used to detect and localize significant low-level features of an image better than traditional gradient operators. Moreover, based on physiological and psychophysical evidences, the PC theory provides a simple but biologically plausible model of how mammalian visual systems detect and identify features in an image [48,49]. Furthermore, PC can be considered as a dimensionless measure for the significance of local structure [50]. So in this paper, we employ PC map to characterize low-level image features.

There are many ways to compute PC. In this paper, the method developed by Kovesi [51] is employed, which was based on a local energy model. To compute the PC of the 2-D image $I(x, y)$, a transfer function of 2-D log-Gabor filter is needed. A 2-D log-Gabor filter is constructed by using Gaussian as the spreading function [52] in several orientations, which is given as follows.

$$G(\omega, \theta_o) = \exp\left(\frac{-\left(\log\left(\omega/\omega_0\right)\right)^2}{2\sigma_r^2}\right) \cdot \exp\left(\frac{-\left(\theta-\theta_o\right)^2}{2\sigma_\theta^2}\right) \tag{1}$$

where $\omega_0$ is the filter's center frequency, $\sigma_r$ controls the filter's bandwidth. $\theta_o = o\pi/O$, $\theta_o = o\pi/O$, $o = \{0, 1, \ldots, O-1\}$ is the orientation angle of the filter, $O$ is the number of orientations and $\sigma_\theta$ is the standard deviation of the Gaussian function in angular orientation, which determines the filter's angular bandwidth. Assuming that $M_{no}^o$ and $M_{ne}^e$ are the odd symmetric and even symmetric components of the 2-D log-Gabor filter (Eq. (1)) at scale $n$ and orientation $o$, and they form a quadrature pair. The response vector at scale $n$ and orientation $o$ is obtained by the convolution of each quadrature pair with the input image $I(x, y)$, and is given as follows.

$$\left[e_{no}(x, y), o_{no}(x, y)\right] = \left[I(x, y) * M_{no}^e, I(x, y) * M_{no}^o\right] \tag{2}$$

The amplitude of the response $A_{no}$ and the phase angle $\varphi_{no}$ at scale $n$ and orientation $o$ are given by as follows.

$$A_{no} = \sqrt{e_{no}^2(x, y) + o_{no}^2(x, y)} \tag{3}$$

(a) original image (b) PC map (c) gradient map

**Fig. 1.** Advantage of PC map over gradients map.

$$\varphi_{no} = \tan^{-1}\left(\frac{o_{no}(x)}{e_{no}(x)}\right) \tag{4}$$

Hence, the PC of an image $I(x, y)$ can be computed over various scales and orientations as follows.

$$PC(x,y) = \frac{\sum_o \sqrt{\left(\sum_n e_{no}(x,y)\right)^2 + \left(\sum_n o_{no}(x,y)\right)^2}}{\varepsilon + \sum_o \sum_n A_{no}(x,y)} \tag{5}$$

where $\varepsilon$ is a small number to avoid division-by-zero. To some extent, PC is sensitive to distortion. This is conducive to IQA. In essence, IQA evaluates how the noise affects human perception. When the noise (distortion) is introduced, low-level features extracted by PC would be changed. For example, Gaussian blur smooths edges and corners, and corresponding values of PC would become lower from relatively high values. Meanwhile, the corresponding LBPs of PC would be changed from one type to another in their own ways, determining by the characteristics of low-level features. Besides, for other types of distortions, the LBPs will be changed accordingly. As a result, exploiting the sensitivity to distortion of PC and extracting LBP in PC domain will benefit IQA.

In the experiments, the number of scales is set as 4 and the number of filter orientations is set as 6 for more low-level features. The range of PC lies between the values 0 and 1. In order to visualize the PC map for friendly-viewing and calculate the LBP in PC map, we normalized the values of PC to range [0,255]. Although PC is contrast invariant, we still utilize PC to represent low-level image features (structure) rather than image gradients. The reason is that PC can capture more fine details of image structure, which are instrumental for improving the accuracy of quality prediction, while image gradients cannot. Fig. 1 shows advantages of PC map over gradient map in capturing finer features. Fig. 1(a) is original image, coming from MLIVE database, Fig. 1(b) is PC map, and Fig. 1(c) is gradient map. As shown in Fig. 1, the PC map has more fine details than gradient map, for example, the roof tiles of the building, tree leaves and mountains in the top are very clear (for better viewing, please zoom in the pictures).

Although PC and GM maps look a bit alike, they have different meanings. In PC, low-level features are represented by moment of PC covariance which includes not only edge strength but also corner strength and other features. PC believed that features are not simple step changes (discontinuity) in luminance [49], while gradient-based methods extract features based on step changes and do not correctly detect and localize features. So PC can extract more features that are finer. Moreover, feature detection and identification in the image represented by PC are more consistent with the HVS perception. Therefore, the characteristics of the PC are instrumental for improving the accuracy of quality prediction. However, PC is contrast-invariant which is not conducive to IQA. GM is contrast sensitive which can contribute to IQA. Therefore, GM can compensate for the contrast invariance of PC. We only use the different characteristics of PC and GM. In the next section, GM is exploited for LBP histogram creation in PC domain.

### 3.2. Image gradient magnitude

The HVS is sensitive to low-level features represented by PC. Such features often attract eye's attention at first sight, and these features can reflect image structure. However, using PC map to characterize image structure has limitations due to its contrast invariance. Therefore, we propose that using image gradient magnitude as complementary features to capture the changes of image contrast. When more than two stimuli are in an image whose response values are greater than eyes' response threshold, the eyes would be attracted to the stimulus which has the largest response value automatically and rapidly due to foveation [53] according to the visual attention principle. In this paper, a more complex image gradient computation method is used, which is the same as in [45].

$$F_1 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 1 & 3 & 8 & 3 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ -1 & -3 & -8 & -3 & -1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad F_2 = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 8 & 3 & 0 & 0 \\ 1 & 3 & 0 & -3 & -1 \\ 0 & 0 & -3 & -8 & 0 \\ 0 & 0 & -1 & 0 & 0 \end{bmatrix} \quad F_3 = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 3 & 8 & 0 \\ -1 & -3 & 0 & 3 & 1 \\ 0 & -8 & -3 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 \end{bmatrix} \quad F_4 = \begin{bmatrix} 0 & 1 & 0 & -1 & 0 \\ 0 & 3 & 0 & -3 & 0 \\ 0 & 8 & 0 & -8 & 0 \\ 0 & 3 & 0 & -3 & 0 \\ 0 & 1 & 0 & -1 & 0 \end{bmatrix}$$

**Fig. 2.** Image gradient filters.

$$G = \max\left(\frac{1}{16} I * F_{n=1,2,3,4}\right) \tag{6}$$

$F_{n=1,2,3,4}$ are four gradient filters, shown in Fig. 2. The star sign (*) denotes the convolution operation. $I$ denotes the image. $G$ denotes the image gradients magnitude map. Here, Sobel or Prewitt filters are not used since these filters are small ($3 \times 3$) and only have horizontal and vertical directions, resulting in insufficient structure information in the neighborhood. Instead, we use 4 directions, larger size filters to obtain a better image gradients map which is more consistent with the HVS.

In order to integrate IQA on image gradient into IQA on distribution of low-level features in a single representation form, image gradients are accumulated on each bin of the LBP histogram in PC domain in the next section.

## 4. The proposed NR IQA method

In this section, we first represent distribution of low-level image features (structure) with local binary pattern (LBP) in PC domain. Then an image gradient magnitude map is employed as a weighting factor to build the LBP histogram to represent the distribution of low-level features (structure), which overcomes the limitation of PC contrast invariance. Finally, gradient-weighted LBP histograms at multi-scale are combined as the input of support vector regression (SVR) to predict quality.

### 4.1. LBP based structure extraction in PC domain

The spatial correlations between the central pixel and its neighborhood are analyzed with the relative intensity relationship. Local binary pattern (LBP) [54] is designed for structure description. And LBP achieves great success in structure description for texture classification. In this paper we employ the LBP to describe the distribution of structure in PC domain. According to the relative intensity relationship between the central pixel$x_c$ and its circularly symmetric neighbor pixels$x_i$, the LBP of $x_c$ is defined as follows.

$$LBP_{P,R}(x_c) = \sum_{i=0}^{P-1} s(I_i - I_c)\, 2^i \tag{7}$$

$$s(I_i - I_c) = \begin{cases} 1, I_i - I_c \geq 0 \\ 0, I_i - I_c < 0 \end{cases} \tag{8}$$

where $I_i$ and $I_c$ are the values of the central pixel $x_c$ and its neighbor $x_i$ in PC map, P is the number of neighbors, and R is the radius of the neighborhood. Meanwhile, Ojala [55] also investigated the uniform

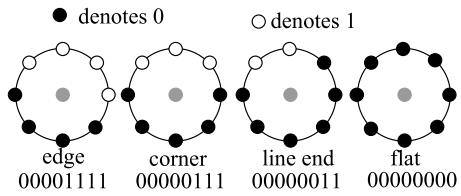


**Fig. 3.** Illustration of changes of LBPs.

LBP patterns, which can provide majority structural information (almost 90%), and the uniform and rotation invariant pattern is defined as follows.

$$LBP^{ur}_{P,R}\left(x_i\right)=\begin{cases}\sum_{i=0}^{P-1} s\left(I_i-I_c\right), & if\,\mu\left(LBP_{P,R}\left(x_i\right)\right)\le 2\\ P+1, & else\end{cases} \tag{9}$$

where $\mu\left(LBP\left(x_i\right)\right)=\left|s\left(I_{P-1}-I_c\right)-s\left(I_0-I_c\right)\right|+\sum_{i=1}^{P-1}\left|s\left(I_i-I_c\right)-s\left(I_{i-1}-I_c\right)\right|$.

With (9), the LBP value of each pixel in PC map can be calculated. For R = 1, P = 8, there are 10 types of LBPs. In the latter experiments, we found that for greater values of P and R, the computational complexity significantly increases, and only make little contribution to the performance of IQA. So in this paper, R and P are set as 1 and 8. After getting PC-based LBPs, a gradient-weighted LBP histogram is built to represent image structure distribution.

With image distortions, the PC-based LBP of a pixel can change from one type to another (as to be elaborated in experimental section). Different PC-based LBPs denote different local structure. Moreover, different kinds of distortions result in different changes to the structure with the accompanying changes to its PC-based LBPs. For example, Gaussian blur mainly degrades edges in the image. As a result, an edge pattern may be distorted into a flat pattern. On the other hand, JPEG compression mainly causes blockiness artifact. As a result, a flat pattern may be distorted into an edge pattern. Specifically, the PC-based LBP '00001111' of a pixel may represent an edge feature, as shown in Fig. 3. When distortions are present, the PC-based LBP pattern would be altered, resulting in '00011111' which represents corner feature or '00111111' which represents line end feature. In summary, based on the analysis above, LBPs in PC domain can effectively capture image degradation due to multiple distortions via describing the spatial correlations between the central pixel and its neighborhood in PC domain.

### 4.2. Gradient-weighted LBP histogram in PC domain at multi-scale

In Sections 2.2 and 3.1, image gradient magnitude and LBP in PC domain are introduced. Due to the limitation of contrast invariance of PC, image gradient features are investigated to compensate for PC. In order to highlight the pixels with severe degradation on contrast while weaken the pixels with minor degradation on contrast, we employ image gradient magnitude as a weighting factor for LBP-based histogram creations in PC domain (i.e., image gradient of each pixel is accumulated on each bin).

$$H^{LBP}(b)=\frac{1}{MN}\sum_{i=0}^{M-1}\sum_{j=0}^{N-1}\omega\left(LBP^{ur}_{P,R}\left(x_{ij}\right),b-1\right),b\in[1,10] \tag{10}$$

$$\omega\left(LBP^{ur}_{P,R}\left(x_{ij}\right),b-1\right)=\begin{cases}G_{ij}, & if(LBP^{ur}_{P,R}(x_{ij}))=b-1\\ 0, & else\end{cases} \tag{11}$$

where $H^{LBP}(b)$ denotes gradient-weighted LBP histogram of distorted images in PC domain. $b$ is the index of bin of the histogram. The number of bins is 10. In such a way, we emphasize image regions with high contrast changes and take both the structural and contrast information in a single representation.

The perceivability of image details depends on the sampling density of the image, the distance from the image plane to the observer and the perceptual capability of the observer's visual system. In practice, the subjective evaluation of images varies when these factors vary [56]. A single-scale gradient-weighted LBP histogram may be appropriate only for specific settings. Multi-scale methods are a convenient way to incorporate image details at different resolutions. Here, we extract gradient-weighted LBP histogram features in PC domain at multiple scales for IQA, and a combined gradient-weighted LBP histogram is defined as follows.

$$H^{LBP}_m=concat_{i=1}^{n\_scale}\left(\left(H^{LBP}_i\right)^{\beta_i}\right) \tag{12}$$

where $H^{LBP}_m$ denotes the combined histogram. n_scale denotes the number of scales and 'concat' operator denotes combining all the LBP histograms in PC domain at multi-scale into a larger histogram. $\beta_i$ is used to adjust the relative importance of different scales, $\beta_1$, $\beta_2$, $\beta_3$, $\beta_4$, $\beta_5$ = 0.0448, 0.2856, 0.3001, 0.2363, 0.1333 [56]. Besides the original image scale, the coarser scale is formed by down sampling by a factor of 2 in each dimension. In the latter experiments, we found that when n_scale is set as 5 can get better results for IQA, resulting in a combined gradient weighted LBP histogram vector to characterize the image structure. The number bin of the combined histogram is 50 in total (i.e., we use 50 features to evaluate the quality of an image).

### 4.3. Image quality assessment

With Eq. (12), we can extract a histogram-based feature vector from a given distorted image with 50 dimensions. However, how to effectively pool feature vectors from distorted images to predict their quality scores with high consistency with the subjective perception is an open problem. Since 50-D histogram-based feature vectors can represent different structure with different degradations on low-level image features (structure), leading to different level quality scores. Moreover, pooling the histogram-based feature vectors for quality score is a crucial step in our method. Support vector regression (SVR) based pooling features method is a good choice, and we learn SVR model using histogram-based vectors extracted from image databases to predict quality scores.

Given training dataset $\{(x_1,y_1),(x_2,y_2),\ldots,(x_r,y_r)\}$, where $x_i$ is the 50-D feature vector and $y_i$ is ground truth (quality scores from image database), the SVR model is employed to map the feature space to quality score. The SVR [57] can be formulated as follows.

$$\min_{\omega,\delta,\upsilon,\upsilon^*}\frac{1}{2}\omega^T\omega+C(\sum_{i=1}^{r}\upsilon_i+\sum_{i}^{r}\upsilon_i^*)$$
$$s.t.\begin{cases}y_i-\omega^T\phi(x_i)-\delta\le\varepsilon+\upsilon_i,\\ \omega^T\phi(x_i)+\delta-y_i\le\varepsilon+\upsilon_i^*,\\ \upsilon_i,\upsilon_i^*\ge 0,i=1,2,\ldots,r.\end{cases} \tag{13}$$

where $\omega$ is the weight vector, $\delta$ is the bias parameter, $\upsilon_i,\upsilon_i^*$ are the slack variables, $C$ is a hyper-parameter. $K(x_i,x_j)=\phi(x_i)^T\phi(x_j)$ is the kernel function. In this paper, the radial basis function (RBF) kernel $K(x_i,x_j)=\exp(\gamma\left\|x_i-x_j\right\|^2)$ is employed, $\gamma$ is the parameter of the kernel. After training SVR model, the output of the regression model is the predicted quality score. In the experiments, we adopt the LibSVM [57] to train the SVR model as follows. The grid search technique with 10-fold cross-validation is employed to tune the hyper-parameters.

$$SVR_{model}=SVR_{train}\left(H^{LBP}_{m,train}(I),Gtruth\right) \tag{14}$$

where $H^{LBP}_{m,train}(I)$ denotes histogram-based feature vectors extracted from the train dataset, Gtruth denotes ground truth of the quality score (DMOS, Difference Mean Opinion Score) and $SVR_{model}$ denotes

a trained model. An image quality score can be calculated using the trained model.

$$Q(I) = SVR_{model}\left(H^{LBP}_{m,test}(I)\right) \tag{15}$$

$H^{LBP}_{m,test}(I)$ denotes histogram-based feature vectors extracted from test dataset. $Q(I)$ denotes the predicted quality score.

## 5. Experimental results and discussions

To quantitatively validate the performance of the proposed method, experiments are conducted with regards to three aspects using benchmark databases. (1) Performance comparison with the state-of-the-art FR IQA methods. (2) Performance comparison with the state-of-the-art NR IQA methods. (3) The computational cost compared with IQA methods. The Matlab source code and validation results can be found at https://github.com/miaoxikui/.

### 5.1. Experiment protocol

In order to make a comprehensive analysis on the performance, we first compare the proposed method with existing FR IQA methods and then compare the proposed method with existing NR IQA methods to demonstrate the performance on two multiply-distorted image databases (MLIVE [13], MDID2013 [14]). MLIVE database has 450 distorted images and consists of two parts. The first part includes 15 reference and 225 contaminated images which are distorted by GB+JPEG (Gaussian Blur followed by JPEG compression). The second part includes the same 15 reference and 225 contaminated images which are distorted by GB+WN (Gaussian Blur followed by white noise). The MDID2013 database has 12 reference and 324 distorted images which are successively corrupted by three types of distortions (GB+JPEG+WN).

The performance of an IQA method is evaluated as the correlation between the predicted scores and the subjective qualities (Difference MOS (DMOS) values). In this paper three criteria are employed: Spearman rank-order correlation coefficient (SRCC) for prediction monotonicity, Pearson linear correlation coefficient (PLCC) and root-mean-squared error (RMSE) for prediction accuracy. The latter two criteria are calculated after the nonlinear logistic mapping function [6] defined below. A good IQA algorithm returns high PLCC/SRCC values and a low RMSE value.

$$q(x) = \beta_1\left(\frac{1}{2} - \frac{1}{1+\exp\left(\beta_2\left(x-\beta_3\right)\right)}\right) + \beta_4 x + \beta_5 \tag{16}$$

where $x$ is the objective quality score and $q(x)$ is the mapped score to the range of subjective scores. $\beta_1, \beta_2, \beta_3, \beta_4, \beta_5$ are five parameters to be fitted.

In the experiments, an 80%/20% training-testing split is used on each database (MLIVE or MDID2013). In order to eliminate the performance bias, training-testing split is repeated 100 times, and the average performance is calculated for the final result.

### 5.2. Performance comparison with FR IQA methods

We compare the proposed method with SSIM [16], GMSD [17], IW-SSIM [58], OSS-SSIM [59], MAD [60], ADM [61], FSIM [62], GMS [63], IGM [64], VSI [65], VIF [66], VSNR [67]. The performance of the proposed method and compared methods (FR) on two databases is listed in Table 1. To avoid mistakes during implementation, the results of MAD is obtained from original author's paper. Others are obtained through demos provided by the authors.

In our proposed method, the number of scales and filter orientations are set as 4 and 6 for low-level features in PC computation. For LBP, uniform and rotation invariant LBP is chosen, and R = 1, P = 8 for high accuracy and efficiency, resulting in 10 bins in the histogram at each scale (i.e. 50 bins in the combined histogram at five scales). SVR model training process is employed on both MLIVE and MDID2013 databases respectively. That is, SVR model is trained and evaluated for each database separately. For training, the grid search technique with 10-fold cross-validation is employed on each training dataset to tune the hyper-parameters $C$, $\gamma$ in Eq. (13) for higher accuracy of quality score prediction.

**Table 1**
Compared performance with FR IQA methods.

| Methods | MLIVE(450 images) | | | MDID2013(324 images) | | |
|---|---|---|---|---|---|---|
| | SRCC | PLCC | RMSE | SRCC | PLCC | RMSE |
| SSIM | 0.902 | 0.926 | 6.969 | 0.622 | 0.656 | 0.037 |
| IW-SSIM | 0.911 | 0.931 | **6.626** | **0.890** | 0.890 | 0.022 |
| OSS-SSIM | **0.919** | **0.931** | **6.681** | 0.763 | 0.729 | 0.033 |
| VIF | **0.915** | **0.932** | 6.761 | **0.905** | **0.915** | **0.020** |
| VSNR | 0.828 | 0.880 | 8.881 | 0.636 | 0.652 | 0.037 |
| MAD | 0.894 | 0.914 | 7.608 | 0.856 | 0.860 | 0.025 |
| ADM | 0.909 | 0.924 | 7.051 | 0.830 | 0.848 | 0.026 |
| FSIM | 0.895 | 0.917 | 7.307 | 0.749 | 0.770 | 0.031 |
| GMS | 0.887 | 0.913 | 7.429 | 0.785 | 0.804 | 0.029 |
| IGM | 0.889 | 0.923 | 7.195 | 0.878 | 0.882 | 0.022 |
| VSI | 0.877 | 0.910 | 7.658 | 0.730 | 0.744 | 0.032 |
| GMSD | 0.880 | 0.911 | 7.656 | 0.877 | **0.902** | **0.021** |
| Proposed | **0.955** | **0.963** | **5.431** | **0.923** | **0.929** | **0.018** |

The best three IQA criteria are highlighted in boldface in Table 1. Most of existing FR IQA methods show poor quality prediction accuracy on both MLIVE and MDID2013, because they are designed for single type of distortions. Whereas, our proposed method outperforms the other FR IQA methods in terms of SRCC, PLCC, RMSE (i.e., the proposed method has greater PLCC, SRCC values and smaller RMSE value than others). Moreover, the proposed method is NR IQA while the other methods need reference image. Overall, the performance of the proposed method on MLIVE is better than that on MDID2013 due to the more complex distortions on MDID2013, and results are in line with the experimental results given by Chandler [15], confirming that multiply distorted databases(MLIVE and MDID2013) challenge most existing FR IQA methods.

### 5.3. Performance comparison with NR IQA methods

**A. Overall Performance on Individual Databases:** In order to demonstrate the effectiveness of the proposed method compared with existing NR IQA methods on the databases, we select both NR IQA methods which are designed for single type of distortions and NR IQA methods which are developed for multiply-distorted images. NFERM [24], GMLOG [26], CORNIA [27], BIQI [28], DIIVINE [29], BLIINDS-II [30], BRISQUE [31], NIQE [32], ILNIQE [34] are designed for single type of distortions. SISBLIM [14], FISBLIM [41], GWH-GLBP [42], LU [43], BQASD [44], IBQASD [45], and SHANIA [46] are developed for multiply-distorted images. Most of the results of the existing NR IQA methods are obtained from the original papers or the source codes downloaded from original authors' website, some results on databases are not listed. The performance of the proposed method and NR IQA methods on MLIVE and MDID2013 is listed in Table 2. The best three IQA criteria are also highlighted in boldface.

BRISQUE, BLIINDS-II, DIIVINE, GMLOG, NFERM, BIQI, SISBLIM, GWH-GLBP, and our proposed method are based on SVR, and the RBF kernel is chosen as their original papers suggested. The grid search technique with 10-fold cross-validation is employed on each training dataset to tune the hyper-parameters. For NIQE, ILNIQE, and CORNIA, we configure them according to the original papers. Due to unavailability of the codes of LU, FISBLIM, SHANIA, BQASD and IBQASD, we take the results from their original or related papers.

In Table 2, IBQASD1 denotes that the corresponding method employing traditional SVR training (one SVR model). IBQASD2 denotes that the corresponding method trained M (M = 5) SVR models and combine these SVRs to construct a more powerful model to solve the

**Table 2**
Compared performance with NR IQA methods.

| IQA method | MLIVE(450 images) | | | MDID2013(324 images) | | |
|---|---|---|---|---|---|---|
| | SRCC | PLCC | RMSE | SRCC | PLCC | RMSE |
| NIQE | 0.789* | 0.858 | 9.489 | 0.614* | 0.645 | 0.037 |
| ILNIQE | 0.900* | 0.914 | 7.538 | 0.707* | 0.709 | 0.034 |
| BLIINDS-II | 0.887* | 0.904 | 7.981 | 0.808* | 0.844 | 0.027 |
| DIIVINE | 0.866* | 0.898 | 8.257 | 0.836* | 0.848 | 0.026 |
| CORNIA | 0.900* | 0.916 | 7.586 | 0.898* | 0.904 | 0.020 |
| BRISQUE | 0.900* | 0.924 | 7.143 | 0.819* | 0.833 | 0.027 |
| GMLOG | 0.833* | 0.872 | 9.164 | 0.824* | 0.830 | 0.027 |
| NFERM | 0.898* | 0.917 | 7.459 | 0.855* | 0.871 | 0.024 |
| BIQI | 0.883* | 0.905 | 7.833 | 0.863* | 0.883 | 0.023 |
| GWH-GLBP | 0.944* | 0.949 | 5.873 | 0.908* | 0.913 | 0.019 |
| FISBLIM | 0.857* | 0.880 | 8.979 | – | – | – |
| SISBLIM | 0.907* | 0.925 | 7.194 | 0.885* | 0.885 | 0.023 |
| BQASD | 0.952* | 0.956 | 5.479 | 0.923* | 0.935 | 0.017 |
| IBQASD1 | **0.952*** | **0.956** | **5.552** | **0.923*** | **0.935** | **0.017** |
| IBQASD2 | **0.958** | **0.960** | **5.445** | **0.929** | **0.940** | **0.017** |
| SHANIA | 0.777* | 0.735 | – | – | – | – |
| LU | 0.908* | 0.942 | – | – | – | – |
| Proposed | **0.955** | **0.963** | **5.431** | **0.923** | **0.929** | **0.018** |

overfitting problem as well as reducing the time complexity. As shown in Table 2, the best NR IQA method on MLIVE is our proposed method, followed by IBQASD2, IBQASD1, BQASD, GWH-GLBP, LU, SISBLIM and BRISQUE. The best NR IQA method on MDID2013 is IBQASD2, followed by IBQASD1, BQASD, our proposed method, GWH-GLBP, CORNIA, SISBLIM and BIQI. The performance of our proposed method is very close to the IBQASD2 (best method). Because IBQASD2 applied the random subspace method in the feature space, and constructed a powerful model using several trained SVRs, this strategy can improve the performance of the method.

It should be noted that the results of some NR IQA methods are not given in Table 2. The same as the results in Table 1, the performance of all the methods on MDID2013 is poorer than that on MLIVE. Most of the NR IQA methods developed for multiply-distorted images outperform the methods designed for single type of distortions, which confirm Chandler's conclusion again [15]. Therefore, it is very urgent to design new NR IQA methods for images distorted by multiple types of distortions.

Furthermore, we evaluate the statistical significance [68] using the T-test with 95% confidence level between SROCC generated by the compared NR IQA methods and our proposed method in 1000 iterations. The results are shown in Table 2, the star sign (*) denotes that our proposed method are statistically better than the compared NR IQA methods. From Table 2, we can see that our proposed method performs significantly better than most of the NR-IQA methods, slightly inferior to IBQASD2 on the two databases. However, our proposed method yielded better results than IBQASD2 in terms of PLCC and RMSE on MLIVE database.

**B. Performance on individual combination of distortions:** To further prove the superiority of our proposed method over the competing NR IQA methods, we evaluate the performance of competing NR IQA methods on individual combination of distortions (i.e., GB+JPEG and GB+WN) in MLIVE. Because only one combination of distortions in MDID2013 (i.e., GB+JPEG+WN), the corresponding results on MDID2013 are listed in Table 2. The SRCC comparison on individual combination of distortions in MLIVE is listed in Table 3. The best three NR IQA methods for each combination of distortions are shown in boldface. It should be noted that similar results can be obtained for PLCC and RMSE. Here we only list SRCC for brevity.

Although GWH-GLBP has lower computational complexity as shown in Table 4, the proposed method found to be significantly better than GWH-GLBP in terms of SRCC, PLCC, and RMSE after the significance analysis. GWH-GLBP only considers image structure feature, which is extracted as the gradient-weighted histogram of LBP calculated in gradient map. Whereas our method not only considers image structure

**Table 3**
SRCC comparison on individual combination of distortions in MLIVE.

| IQA method | GB+JPEG | GB+WN | IQA method | GB+JPEG | GB+WN |
|---|---|---|---|---|---|
| NIQE | 0.899 | 0.833 | GWH-GLBP | 0.948 | 0.903 |
| ILNIQE | 0.899 | 0.890 | FISBLIM | 0.858 | 0.855 |
| BLIINDS-II | 0.892 | 0.884 | SISBLIM | 0.874 | 0.880 |
| DIIVINE | 0.864 | 0.877 | BQASD | 0.954 | 0.948 |
| CORNIA | 0.904 | 0.900 | IBQASD1 | **0.954** | **0.949** |
| BRISQUE | 0.905 | 0.900 | IBQASD2 | **0.960** | **0.952** |
| GMLOG | 0.865 | 0.817 | SHANIA | 0.801 | 0.753 |
| NFERM | 0.919 | 0.887 | LU | 0.907 | 0.904 |
| BIQI | 0.881 | 0.883 | Proposed | **0.956** | **0.952** |

information, but also takes low-level features and their distribution into account. The proposed method employs image gradient to represent structure feature, uses PC to describe low-level feature, and extracts LBP histogram in PC domain at five scales to characterize the distribution of low-level feature. Then the LBP histogram is weighted by gradient magnitude, fusing structure feature and the distribution of low-level feature into a single representation.

According to PC theory [47,48], phase of an image conveys more structural information than gradient. PC of an image can capture more fine details of structure than gradient as illustrated in Fig. 1. The results presented in Tables 2 and 3 show that LBP histogram in PC domain (the proposed method) is better than that taking LBP histogram in gradient domain (as done in GWH-GLBP) for IQA. Because human eyes are sensitive to low-level image features for perception and PC provides a biological and psychophysical model to interpret the detection and identification of low-level feature, making feature perception more consistent with the HVS. The proposed method would yield better performance due to employing more features which are consistent the HVS perception [68]. Therefore, considering both gradient and PC would improve the performance of IQA.

IBQASD1 and IBQASD2 have lower computational complexity shown in Table 4 and very similar performance with our proposed method. They extract the first- and high-order structural features from gradient-magnitude and contrast-normalized maps. In order to avoid overfitting, random subspace method is employed [45] in IBQASD2. It performs the bootstrapping in the feature space. After random sampling, IBQASD2 generates a small subset of features from feature space to reduce the discrepancy between the training data size and the feature vector length. Although IBQASD1 and IBQASD2 employ first- and high-order features to characterize structure and texture information for improving accuracy of IQA, our proposed method (a SVR is employed) outperforms IBQASD1(a SVR is employed) in terms of overall performance on MLIVE shown in Table 2 and performance on individual combination of distortions in MLIVE shown in Table 3. Although 5 SVRs are trained using the subset of features to improve the accuracy of IQA, our proposed method yields better results than IBQASD2 on MLIVE and performs almost the same as IBQASD2 on MDID2013.

Comprehensive experiment results show that the PC map can effectively represent low-level fine features (structure), which is consistent with detection and identification of features in images biologically. LBP map in PC domain can sensitively capture the changes of image structure due to degradation by multiple distortions. Gradient-weighted LBP histogram in PC domain at multi-scale can measure the changes of contrast information and structure to which the HVS is highly sensitive, compensating for the limitation of PC contrast invariance. SVR procedure is utilized to pool histogram features for quality score. These are beneficial for improving the accuracy of quality prediction for multiply-distorted images. Therefore, the quality scores produced by our proposed method are highly correlated with the subjective scores.

**Table 4**
Time cost of each compared IQA method.

| | | | | | | |
|---|---|---|---|---|---|---|
| IQA method | SSIM | IW-SSIM | OSS-SSIM | VIF | VSNR | ADM |
| runtime (s) | 0.1385 | 0.9812 | 0.9815 | 1.3761 | 0.7895 | 0.7893 |
| IQA method | **FSIM** | **GMS** | **IGM** | **VSI** | **GMSD** | **NIQE** |
| runtime (s) | 0.7510 | 0.1529 | 3.4091 | 0.6752 | 0.0113 | 0.4596 |
| IQA method | **ILNIQE** | **BLIINDS-II** | **DIIVINE** | **CORNIA** | **BRISQUE** | **GMLOG** |
| runtime (s) | 6.9016 | 102.7232 | 21.3212 | 3.6845 | 0.5612 | 1.3692 |
| IQA method | **NFERM** | **BIQI** | **GWH-GLBP** | **SISBLIM** | **BQASD** | **IBQASD1** |
| runtime (s) | 3.7959 | 2.0191 | 1.7101 | 2.0762 | 2.3542 | 2.3683 |
| IQA method | **IBQASD2** | **Proposed** | | | | |
| runtime (s) | 2.8542 | **3.2102** | | | | |

### 5.4. Computational cost

The computational cost of each compared IQA method was also evaluated. Experiments were performed on a HP desktop (Intel Core i5-2400 CPU@3.1 GHz and 6G RAM). The software platform was Matlab R2018a. The time cost (unit:s) consumed by each method is average out by 100 distorted images (1280 × 720 from MDID2013) and is listed in Table 4. As for the compared FR IQA methods, we directly run the code to get run-time except MAD due to unavailability of its code. As for OSS-SSIM, we only get p-file code of Matlab (it has been optimized). For the NR IQA methods, BRISQUE, BLIINDS-II, DIIVINE, GMLOG, NFERM, BIQI, SISBLIM, GWH-GLBP, BQASD, IBQASD1, IBQASD2, and our method are based on SVR. The RBF kernel is chosen as their original papers suggested, and the grid search technique is also employed. After training, we track the run-time of quality prediction of the 9 methods. As for NIQE, ILNIQE, and CORNIA, we directly apply the code of these three methods on the images according to the corresponding papers and track the run-time of quality prediction. Due to unavailability of the code of MAD, LU, FISBLIM, and SHANIA, the run-time of these four methods is not listed. As shown in Table 4, our proposed method has a moderate computational complexity.

### 5.5. Discussions

The proposed method obtains promising performance in perceptual quality prediction. The performance improvement is mainly from the following aspects. Firstly, PC map can effectively represent low-level image features, which is highly consistent with the detection and identification features in images of the HVS. Moreover, PC map, to some extent, can characterize image structure to which the HVS is very sensitive and contributes to IQA. Secondly, since the HVS is sensitive to contrast changes, image gradient magnitudes are employed to measure the contrast information at each pixel as complementary features to compensate for the contrast invariance of PC. Meanwhile, gradient magnitudes are used as weighting factors for LBP histogram creation in PC domain, which can highlight image regions with high contrast changes. Thirdly, the gradient-weighted LBP histograms in PC domain at multiple scales are combined by scale factors which are used to adjust the relative importance of different scales to reflect the multi-scale characteristics of the HVS. SVR can effectively map the multi-scale gradient-weighted LBP histogram features into the visual quality.

## 6. Conclusions

In image communication systems, images are usually distorted by multiple types of distortions. However, most existing NR IQA methods mainly focus on single type of distortion. To address this, we proposed a new NR IQA method based on low-level features and their distribution for multiply-distorted images. The underlying principle of proposed method is that distortions not only destroy low-level image features, but also distort the distribution of these features. In order to represent low-level image features, phase congruency (PC) is investigated and LBPs based on PC map are used for describe the distribution of low-level features. To handle the contrast invariance of PC, image gradient magnitudes (GM) are employed as weighting factors to build LBP histogram of the distorted image. PC and GM are complementary and they can reflect different aspects of the HVS for IQA. And then, quality score is predicted by learning a support vector regression model using multi-scale gradient weighted LBP histogram feature vectors. Experimental results on benchmark databases have demonstrated that the proposed method achieves high consistency with subjective perception and performs better than other state-of-the-art methods.

## Acknowledgments

This work is supported by the Robotic Vision Lab (RVL) at Brigham Young University, USA and China Scholarship Council (CSC) program under Grant 201600930087.

## References

[1] L. Liu, Y. Hua, Q. Zhao, H. Huang, A. Bovik, Blind image quality assessment by relative gradient statistics and adaboosting neural network, Signal Process., Image Commun. 40 (1) (2016) 1–15.
[2] C. Bampis, Z. Li, A.K. Moorthy, I. Katsavounidis, A. Aaron, A. Bovik, Study of temporal effects on subjective video quality of experience, IEEE Trans. Image Process. 26 (11) (2017) 5217–5231.
[3] S. Athar, A. Rehman, Z. Wang, Quality assessment of images undergoing multiple distortions stages, in: IEEE ICIP, Beijing, China, 2017.
[4] K. Gu, J. Zhou, J.F. Qiao, G.T. Zhai, W. Lin, A. Bovik, No-reference quality assessment of screen content pictures, IEEE Trans. Image Process. 26 (8) (2017) 4005–4018.
[5] W. Lin, C.-C.J. Kuo, Perceptual visual quality metrics: A survey, J. Vis. Commun. Image Represent. 22 (4) (2011) 297–312.
[6] H. Sheikh, M. Sabir, A. Bovik, A statistical evaluation of recent full reference image quality assessment algorithms, IEEE Trans. Image Process. 15 (11) (2006) 3440–3451.
[7] Q. Li, W. Lin, J. Xu, Blind image quality assessment using statistical structural and luminance features, IEEE Trans. Multimed. 18 (12) (2016) 2457–2469.
[8] Q. Li, W. Lin, Y. Fang, BSD: Blind image quality assessment based on structural degradation, Neurocomputing 236 (2017) 93–103.
[9] H.R. Wu, A. Reibman, W. Lin, F. Pereira, S.S. Hemami, Perceptual visual signal compression and transmission, Proc. IEEE 101 (9) (2013).
[10] N. Ponomarenko, L. Jin, O. Ieremeiev, Image database TID2013: Peculiarities, results and perspectives, Signal Process.: Image Commun. 30 (2015) (2013) 57–77.
[11] H.R. Sheikh, Z. Wang, L. Cormack, A. Bovik, LIVE Image Quality Assessment Database Release 2. http://live.ece.utexas.edu/research/quality (Accessed 21.05.18).
[12] E.C. Larson, D.M. Chandler, Categorical Image Quality (CSIQ) Database. http://vision.okstate.edu/csiq (Accessed 21.05.18).
[13] D. Jayaraman, A. Mittal, A. Moorthy, Objective quality assessment of multiply distorted images, in: Proc. of Int. Conf. Rec. 46th Asilomar Conf. Signals Syst. Comput., 2012, pp. 1693–1697.
[14] K. Gu, G. Zhai, X. Yang, Hybrid no-reference quality metric for singly and multiply distorted images, IEEE Trans. Broadcast. 60 (3) (2014) 555–567.
[15] D.M. Chandler, Seven challenges in image quality assessment: Past, present, and future research, ISRN Signal Process. (2013) 1–53.
[16] Z. Wang, A. Bovik, H.R. Sheikh, Image quality assessment: From error visibility to structural similarity, IEEE Trans. Image Process. 13 (4) (2004) 600–612.
[17] W. Xue, L. Zhang, X. Mou, Gradient magnitude similarity deviation: A highly efficient perceptual image quality index, IEEE Trans. Image Process. 23 (2) (2014) 684–695.
[18] N. Skarbnik, Y. Yehoshua, Zeevi, C. Sagiv, The Importance of Phase in Image Processing. Irwin and Joan Jacobs Center for Communication and Information Technologies (CCIT), Report #773, August, 2010.
[19] J. Wu, M. Zhang, G. Shi, No-reference image quality assessment with orientation selectivity mechanism, in: IEEE ICIP, Beijing, China, 2017.
[20] Y. Fang, J. Yan, L. Li, No reference quality assessment for screen content images with both local and global feature representation, IEEE Trans. Image Process. 27 (4) (2018) 1600–1610.
[21] G. Zhai, X. Wu, X. Yang, A psychovisual quality metric in free-energy principle, IEEE Trans. Image Process. 21 (1) (2012) 41–52.
[22] K. Gu, G. Zhai, X. Yang, No-reference image quality assessment metric by combining free energy theory and structural degradation model, in: Proc. IEEE Int. Conf. Multimedia Expo (ICME), San Jose, CA, USA, 2013.
[23] K. Gu, G. Zhai, X. Yang, A new reduced-reference image quality assessment using structural degradation model, in: Proc. IEEE Int. Symp. Circuits Syst., Beijing, China, 2013.

[24] K. Gu, G. Zhai, X. Yang, W. Zhang, Using free energy principle for blind image quality assessment, IEEE Trans. Multimedia 17 (1) (2015) 50–63.
[25] F. Xie, Y. Lu, A. Bovik, Application-driven no reference quality assessment for dermoscopy images with multiple distortions, IEEE Trans. Biomed. Eng. 63 (6) (2016) 1248–1256.
[26] W. Xue, X. Mou, L. Zhang, A. Bovik, Blind image quality assessment using joint statistics of gradient magnitude and laplacian features, IEEE Trans. Image Process. 23 (11) (2014) 4850–4862.
[27] P. Ye, J. Kumar, L. Kang, Unsupervised feature learning framework for no-reference image quality assessment, in: Proc. of IEEE Conf. Comput. Vis. Pattern Recognit., 2012.
[28] A.K. Moorthy, A. Bovik, A two-step framework for constructing blind image quality indices, IEEE Signal Process. Lett. 17 (5) (2010) 513–516.
[29] A. Moorthy, A. Bovik, Blind image quality assessment: from natural scene statistics to perceptual quality, IEEE Trans. Image Process. 20 (12) (2011) 3350–3364.
[30] M.A. Saad, A. Bovik, Blind image quality assessment: A natural scene statistics approach in the DCT domain, IEEE Trans. Image Process. 21 (8) (2012) 3339–3352.
[31] A. Mittal, A.K. Moorthy, A. Bovik, No-reference image quality assessment in the spatial domain, IEEE Trans. Image Process. 21 (12) (2012) 4695–4708.
[32] A. Mittal, R. Soundararajan, A. Bovik, Making a completely blind image quality analyzer, IEEE Signal Process. Lett. 20 (3) (2013) 209–212.
[33] Y. Zhang, A.K. Moorthy, D.M. Chandler, C-DIIVINE: No-reference image quality assessment based on local magnitude and phase statistics of natural scenes, Signal Process., Image Commun. 29 (4) (2014) 725–747.
[34] L. Zhang, L. Zhang, A. Bovik, A feature-enriched completely blind image quality evaluator, IEEE Trans. Image Process. 24 (8) (2015) 2579–2591.
[35] L. Li, Y. Zhou, W. Lin, J. Wu, No-reference quality assessment of deblocked images, Neurocomputing 177 (2016) 572–584.
[36] L. Li, W. Shen, K. Gu, No-reference quality assessment of enhanced images, China Commun. 13 (9) (2016) 121–130.
[37] L. Li, Y. Yan, Z. Lu, J. Wu, No-reference quality assessment of deblurred images based on natural scene statistics, IEEE Access 5 (2017) 2163–2171.
[38] Y. Fang, K. Ma, Z. Wang, No-reference quality assessment for contrast-distorted images based on natural scene statistics, IEEE Signal Process. Lett. 22 (7) (2014) 838–842.
[39] K. Gu, W. Lin, G. Zhai, X. Yang, No-reference quality metric of contrast-distorted images based on information maximization, IEEE Trans. Cybern. 47 (12) (2017) 4559–4564.
[40] L. Li, W. Lin, X. Wang, G. Yang, Khosro Bahrami, Alex C. Kot, No-reference image blur assessment based on discrete orthogonal moments, IEEE Trans. Cybern. 46 (1) (2016) 39–50.
[41] K. Gu, G. Zhai, M. Liu, FISBLIM: A five-step blind metric for quality assessment of multiply distorted images, in: IEEE Workshop on Signal Process. Syst., 2013, pp. 241–246.
[42] Q. Li, W. Lin, Y. Fang, No-reference quality assessment for multiply-distorted images in gradient domain, IEEE Signal Process. Lett. 23 (4) (2016) 541–545.
[43] Y. Lu, F. Xie, T. Liu, No reference quality assessment for multiply-distorted images based on an improved bag-of-words model, IEEE Signal Process. Lett. 22 (10) (2015) 1811–1815.
[44] T. Dai, K. Gu, Z. Xu, Blind quality assessment of multiply-distorted images based on structural degradation, in: IEEE International Conference on Image Processing (ICIP), Vol. 27, Athens, Greece, 2017.
[45] T. Dai, K. Gu, L. Niu, Referenceless quality metric of multiply-distorted images based on structural degradation, Neurocomputing 290 (2018) 185–195.
[46] Y. Li, L.M. Po, X. Xu, No-reference image quality assessment using statistical characterization in the shearlet domain, Signal Process.: Image Commun. 29 (7) (2014) 748–759.
[47] A.V. Oppenheim, J.S. Lim, The importance of phase in signals, Proc. of IEEE 69 (5) (1981) 529–541.
[48] M.C. Morrone, J. Ross, D.C. Burr, R. Owens, Mach bands are phase dependent, Nature 324 (6049) (1986) 250–253.
[49] M.C. Morrone, R.A. Owens, Feature detection from local energy, Pattern Recognit. Lett. 6 (5) (1987) 303–313.
[50] L. Henriksson, A. Hyvärinen, S. Vanni, Representation of cross-frequency spatial phase relationships in human visual cortex, J. Neurosci. 29 (45) (2009) 14342–14351.
[51] P.D. Kovesi, Phase congruency: A low-level image invariant, Psychol. Res.-Psychol. Forschung. 64 (2) (2000) 136–148.
[52] W. Wang, J. Li, F. Huang, Design and implementation of log-gabor filter in fingerprint image enhancement, Pattern Recognit. Lett. 29 (3) (2008) 301–308.
[53] Y. Fang, W. Lin, B. Lee, C. Tong Lau, Z. Chen, C. Lin, Bottom-up saliency detection model based on human visual sensitivity and amplitude spectrum, IEEE Trans. Multimed. 14 (1) (2012) 187–198.
[54] T. Ojala, K. Valkealahti, E. Oja, M. Pietikäinen, Texture discrimination with multidimensional distributions of signed gray-level differences, Pattern Recognit. 34 (3) (2001) 727–739.
[55] T. Ojala, M. Pietikäinen, T. Mäenpää, Multiresolution gray-scale and rotation invariant texture classification with local binary patterns, IEEE Trans. Pattern Anal. Mach. Intell. 24 (7) (2002) 971–987.
[56] Z. Wang, E. Simoncelli, A. Bovik, Multiscale structural similarity for image quality assessment, in: Proc. of 37th Asilomar Conf. on Sig., Sys. and Com., Vol. 2, 2003, pp. 1398–1402.
[57] C.C. Chang, C.-J. Lin, LIBSVM: A library for support vector machines, ACM Trans. Intell. Syst. Technol. (TIST) 2 (2011) 27:1–27:27.
[58] Z. Wang, Q. Li, Information content weighting for perceptual image quality assessment, IEEE Trans. Image Process. 20 (5) (2011) 1185–1198.
[59] K. Gu, M. Liu, G. Zhai, Quality assessment considering viewing distance and image resolution, IEEE Trans. Broadcast 61 (3) (2015) 520–531.
[60] E.C. Larson, D.M. Chandler, Most apparent distortion: Full reference image quality assessment and the role of strategy, J. Electron. Imaging 19 (1) (2010) 011 006-1–011 006-21.
[61] S. Li, F. Zhang, L. Ma, Image quality assessment by separately evaluating detail losses and additive impairments, IEEE Trans. Multimedia 13 (5) (2011) 935–949.
[62] L. Zhang, D. Zhang, X. Mou, FSIM: A feature similarity index for image quality assessment, IEEE Trans. Image Process. 20 (8) (2011) 2378–2386.
[63] A. Liu, W. Lin, M. Narwaria, Image quality assessment based on gradient similarity, IEEE Trans. on Image Process. 21 (4) (2012) 1500–1512.
[64] J. Wu, W. Lin, G. Shi, A. Liu, Perceptual quality metric with internal generative mechanism, IEEE Trans. Image Process. 22 (1) (2013) 43–54.
[65] L. Zhang, Y. Shen, H. Li, VSI: A visual saliency-induced index for perceptual image quality assessment, IEEE Trans. on Image Process. 23 (10) (2014) 4270–4281.
[66] H. Sheikh, A. Bovik, Image information and visual quality, IEEE Trans. Image Process. 15 (2) (2006) 430–444.66.
[67] D. Chandler, S. Hemami, VSNR: A wavelet-based visual signal to-noise ratio for natural images, IEEE Trans. Image Process. 16 (9) (2007) 2284–2298.
[68] ITU-T Recommendation P.1401, Methods, metrics and procedures for statistical evaluation, qualification and comparison of objective quality prediction models, International Telecommunication Union, Geneva, CH, 2012.