Analysis of Structural Characteristics for Quality Assessment of Multiply Distorted Images

Guanghui Yue, Chunping Hou, Ke Gu, Nam Ling, *Fellow, IEEE*, and Beichen Li

Abstract—Perceptual image quality assessment (IQA) plays an important role in numerous applications, including image restoration, compression, enhancement, and others. Although many works have been conducted on individually distorted IQA problems and have achieved encouraging results, few studies have been conducted on multiple distorted (MD) IQA problems. Thus, limited progress has been made. In this paper, we propose a novel no reference image quality assessment (NR-IQA) method, named *improved multi-scale local binary pattern* (IMLBP), for addressing multiply distorted IQA problems. The image structures are sensitive to image distortions, which motivates us to utilize the structural characteristics for overall image quality prediction. We improved the local binary pattern (LBP) by considering the human visual mechanism to better extract the structural information. The IMLBP contains two parts, the LBP and the radius difference LBP (DLBP). The DLBP reflects the values' changes in the radial direction. Specifically, when the radius value is small, the proposed descriptor is computed to represent microstructural information. Conversely, it represents macrostructural information when the radius becomes large. Moreover, to better mimic the human visual mechanism, the IMLBP is computed with the multi-scale strategy and the operation is based on a patch unit whose size is proportional to the radius value. The frequency histogram of feature maps is transformed to feature vectors. Subsequently, a predictable function trained by the support vector regression (SVR) is used to infer the overall quality score. Experimental results show that the proposed method outperforms most state-of-the-art IQA metrics on publicly available multiply distorted image databases.

Index Terms—Image quality assessment (IQA), multiple distortions, local binary pattern (LBP), no reference (NR).

I. INTRODUCTION

WITH the explosive development of visual acquisition, display and processing technologies, digital images have become ubiquitous in our life. However, they may suffer multiple distortions during their acquisition, processing and transmission [1]–[3]. IQA, as the benchmark for evaluating the degree of quality degradation, has monitored and improved the quality of consumer products and networking infrastructures. Based on the availability of human beings, IQA is divided into subjective and objective categories. Objective assessment has more advantages than subjective assessment, which is accurate but time consuming and expensive. To date, several existing works have been conducted to develop a universal objective assessment metric that takes us a step closer to the ultimate perceptual experience of human observers [4]–[8].

Generally, according to the availability of reference information, objective IQA metrics can be categorized into three types: full reference (FR), reduced reference (RR) and no reference (NR) metrics. Since the reference information is usually not available, IQA metrics that do not rely on reference information (e.g., NR-IQA) will be the mainstream direction of future research. Existing general-purpose NR-IQA algorithms are always based on the hypothesis that the extracted features should effectively represent the distorted type and degree. Hence, feature extraction is the key issue for building a high-performance NR-IQA algorithm. According to the extracted features, previous metrics usually follow one of the subsequent two categories: natural scene statistic (NSS)-based approaches and learning-based approaches [9]–[11]. Those NSS-based approaches assume that natural scenes possess certain statistical properties. Statistical models (such as general Gaussian distributions) are used to the characterize statistical properties of cosines [12], contourlet coefficients [13] and wavelets [14]. To some extent, the variation of the model's parameters reflects the degree of introduced distortion. The learning-based approaches depend on a large number of quality-aware features. Then, NR-IQA problem can be attributed as a regression problem via neural network learning. Various regression techniques, such as the SVR [15] and the random forest [16], were applied to learn the mapping from the feature space to the image quality space.

Currently, great achievement has been made in single distorted IQA, while MD IQA has remained in the initial stages. Since an image usually suffers from multiple distortions in practical applications, it is more meaningful to propose an effective MD IQA method. In [17], features that are sensitive to each distortion type were first selected. Next, the correlation was analyzed between the feature and distortion type for better feature extraction and selection. Then, codebooks for different distortion types were constructed to form the bag of words (BoW), which was regarded as the representation of the image. Finally, the image quality can be predicted by the linear combination of each dimension of the BoW. Similarity, Li *et al*. [18] extracted quantities of features from four categorically distorted images. However, a physical combination of individually distorted features may not well explain the final quality since the individual distortion may have joint effects [19]. Inspired by it, Gu *et al*. [20] successfully analyzed

G. Yue, C. Hou, and B. Li are with the School of Electrical and Information Engineering, Tianjin University, Tianjin 300072, China (e-mail: yueguanghui@tju.edu.cn; hcp@tju.edu.cn; relidin@126.com).

K. Gu is with 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).

N. Ling is with the Santa Clara University, Santa Clara CA 95053 USA. (e-mail: nling@scu.edu).

the noise strength, blur and JPEG effects, before the free energy term was added in order to consider the possible interaction of different distortions. In [21], Li *et al.* extracted the LBP features in the gradient domain and used the gradient amplitude as the weight to constitute the feature vectors as the input of the SVR model.

In this paper, we propose a novel objective NR-IQA metric for multiple distorted images without any prior knowledge of distortion combination in an image. The major technical contributions of this work are summarized here. First, motivated by the human visual mechanism, the LBP descriptor is improved to better represent the structural information. Especially, instead of individual pixel gray values at sample points, the image patches centered at the sampling locations are used during the LBP map computation to avoid information loss. Second, the improved LBP descriptor is multi-scale based on reflecting both the microstructural and macrostructural information. The LBP map with a small radius is extracted to encode the image's primitive microstructure. Meanwhile, the DLBP map (which reflects the gray values' change in radial direction) together with the LBP map with a large radius are extracted to encode the macrostructure. Third, the frequency histograms of the LBP and DLBP maps are encoded to form the concatenated histogram, which is transformed as the input of the SVR model. The experimental results on public multiple distortion databases demonstrate that the proposed method outperforms the state-of-the-art methods in the MD IQA problem.

The remainder of this paper is organized as follows. Section II describes the LBP algorithm and its applications in the NR-IQA. The proposed method is fully introduced in Section III. The experimental results and discussions are presented in Section IV. Finally, the conclusion is given in Section V.

II. RELATED WORKS

A. LBP Descriptor

The LBP descriptor [22] has been widely used in facial recognition [23] and texture classification [24]. Since quality degradation usually distorts the image structure, the LBP that describes the local structure by encoding the image into matrix, can be applied to the IQA problem. It describes the relationship between the center pixel n_c and its surrounding neighbors by computing the gray-level difference. It can be expressed as:

$$LBP_{P,R} = \sum_{i=0}^{P-1} s(I(n_i), I(n_c)) \cdot 2^i \qquad (1)$$

where P is the number of neighbors, which controls the quantization of the angular space. R is the radius of the neighbors that determines the spatial resolution, as shown in Fig. 1(a). $I(n_c)$ denotes the gray value of the center pixel n_c, while $I(n_i)$, $i = \{0, 1, 2, 3, \cdots, P\text{-}1\}$ denotes the gray value of circularly symmetric neighbor pixel n_i. The binomial factor 2^i is used to encode each sampling location. The thresholding function $s(\cdot, \cdot)$ is defined as:

$$s(I(n_i), I(n_c)) = \begin{cases} 1, & if \quad I(n_i) - I(n_c) \geq T \\ 0, & if \quad I(n_i) - I(n_c) < T \end{cases} \qquad (2)$$

The threshold T is usually set as zero. By definition, the LBP only has gray invariance and does not have rotational invariance with uniform patterns. Fortunately, this problem can be tackled by the uniformity measure [22]. Then, the rotational invariant uniform LBP can be formulated as:

$$LBP_{P,R}^{riu2} = \begin{cases} \sum_{i=0}^{P-1} s(I(n_i), I(n_c)) & , if \; \mu(LBP_{P,R}) \leq 2 \\ P+1 & , otherwise \end{cases} \qquad (3)$$

where the superscript *riu2* denotes the rotational invariant uniform patterns when the uniform measure μ is less than 2. Formally, μ is calculated as the number of bitwise transitions:

$$\begin{aligned} \mu(LBP_{P,R}) = & \|s(I(n_{P-1}), I(n_c)) - s(I(n_0), I(n_c))\| \\ & + \sum_{i=1}^{P-1} \|s(I(n_i), I(n_c)) - s(I(n_{i-1}), I(n_c))\| \end{aligned} \qquad (4)$$

From Eqs. (3)-(4), it can be intuitively observed that the uniform LBP map would have *P+2* patterns.

B. LBP Applications in NR-IQA

As the distortion may induce the loss of the structural and textural information, it is crucial to unearth and exploit the features that reflect the image's distortion. Currently, many descriptors (such as SIFT, Harris and LBP) have been reported for image feature extraction. Compared to LBP, SIFT and Harris aim to detect local key points, whose positions and numbers vary according to different image contents. For a distorted image and its reference, their detected key points may be different, which must increase the difficulty in designing IQA models. Besides, key points only contain the information in the special location rather the whole information of an image, and this usually makes the IQA model overlook crucial information. Therefore, SIFT and Harris may be not well suitable for IQA. On the contrary, LBP does not rely on key points and can represent image characteristics well by numerical statistics of all pixels. To date, several works have tried to integrate the LBP descriptor into IQA metrics [21], [25]–[27].

In [25], the features extracted from LBP were regarded as complementary factors to improve the final performance. The structural degradation was computed by measuring the change of each bin in the histograms between the reference and distorted images. Similarly, Zhang *et al.* [26] first decomposed the image into multiple-scales using the Laplacian of Gaussian (LoG) filters. Then, each sub-band image was encoded using the proposed generative LBP and concentrated the joint histograms finally. Although the LBP descriptor has made remarkable progress in the extraction of local structure information, it still has inherent drawbacks. For one, the high-order information is ignored since it only encodes the binary result of the first-order derivative among the neighbors. Moreover, LBP is incapable of describing the directional information. Considering such shortcomings, Du *et al.* [27] took the LBP as the tool of first-order derivative information and extracted more high-order local derivative patterns to form the final feature vectors. The results demonstrated that the experimental performance was greatly improved owing to the

complement of high-order information. To further reduce the computational complexity and access better performance, Li *et al*. [21] proposed a novel structural extractor named the GWH-GLBP. First, it calculated the gradient map. Then, the gradient LBP map was accumulated using the gradient-weight of the histogram to better represent the structural and contrast information. Finally, the SVR was adopted to map the feature space to a quality score. The experimental results showed that GWH-GLBP achieved encouraging results on the MD IQA. In summary, the above works validate that the LBP can definitely prompt IQA studies.

III. THE PROPOSED NR-IQA METHOD

This section includes two subsections: 1) multi-scale improved LBP descriptor and its application to IQA; 2) image quality estimation. Specifically, the multi-scale improved descriptor is composed of the LBP descriptor and DLBP descriptor with multiple scales. Moreover, each operation is conducted based on a patch unit, whose size is proportional to the distance between the sampling point and the centered point in order to better mimic the visual mechanism.

A. Multi-scale Improved LBP Descriptor and Its Application to IQA

Fig. 1(a) is the diagram of the LBP descriptor. By definition, the LBP describes the relationship between the center point and its surrounding neighbors and only reflects the microstructure of the image when the radius is small. With the radius increasing, it can reflect macrostructural information to some extent. However, the distance between adjacent sampling points becomes larger as the radius increases if the LBP calculation is still pixel based. As a result, the performance of feature maps might be reduced since most pixels are not included in the computation, leading to information loss. In this paper, we improve the LBP by mimicking the human visual mechanism.
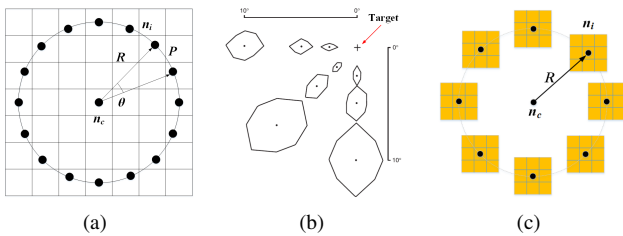

Fig. 1. (a) Standard LBP descriptor. (b) Critical spacing. (Figure adapted from ref. [28]). (c) Improved LBP descriptor.

Human eyes capture the information of both the fixated individual point and its peripheral points when gazing at a point [28]. To be specific, a target may be strongly degraded by the simultaneous presentation of flanking objects. To allow for the unimpaired perception of the target, the distance between the fixated target and the flanking objects is defined as the critical spacing [28]. The critical spacing (which may reflect the minimum area of visual attention) grows in proportion to the object's distance from the fixated target [29]. The sizes

of the interaction regions linearly varies with eccentricity and spatial interaction zones appear to be radially elongated with elliptical shapes, as shown in Fig. 1(b). Inspired by these, we modify the LBP descriptor by replacing the individual pixel gray value at the sampling point with a simple filtered response that is derived from source image patches centered on the sampling location, as shown in Fig. 1(c). Since the interaction space is elliptical, we simply utilize the square patch to reflect the shape characteristic. To be more specific, the $I(n_i)$ is replaced by the filtered response of square patch centered at n_i in the LBP calculation. Formally, given a center pixel n_c and a patch filter Φ, the rotational invariant LBP descriptor is redefined as:

$$LBP_{P,R}^{riu2} = \begin{cases} \sum_{i=0}^{P-1} s(\Phi(N_{i,R,\omega_R}), n_c) & ,if \ \mu(LBP_{P,R}) \leq 2 \\ P+1 & , otherwise \end{cases} \quad (5)$$

where N_{i,R,ω_R}, $i = \{0, 1, 2, 3, \ldots, P\text{-}1\}$ denotes the circularly symmetric local patch with size $\omega_R \times \omega_R$ centered at the neighboring sampling point n_i with distance R (i.e., the radius value) from the centered pixel n_c. In this paper, we follow [21] and set P as 8. $\Phi(N_{i,R,\omega_R})$ is the filtered response of patch N_{i,R,ω_R} centered at n_i. Since the distance between the adjacent sampling points increases as the radius increases, the patch size should be changed with respect to the radius's value change. In this paper, the patch size is proportional to the radius since the shape of the interaction space is radially elongated [30]. We empirically set $\omega_R = R - 1$ and use a median filter. Then, the uniform measure, μ, is redefined as:

$$\mu(LBP_{P,R}) = \|s(\Phi(N_{P-1,R,\omega_R}), I(n_c)) - s(\Phi(N_{0,R,\omega_R}), I(n_c))\| + \sum_{i=1}^{P-1} \|s(\Phi(N_{i,R,\omega_R}), I(n_c)) - s(\Phi(N_{i-1,R,\omega_R}), I(n_c))\|.$$

$$(6)$$

In addition, as the density of photopic retinal photoreceptors rapidly decreases with the distance away from the fovea, the discrimination ability falls off rapidly [31], [32]. Therefore, we also pay more attention to the relationship between different gray values in the radial directions. To mimic this characteristic, we employ the DLBP descriptor. To specify, unlike the LBP descriptor calculating the difference value between the sampling point and center point, the DLBP descriptor computes the difference value between sampling points in the radial directions. Considering the visual mechanism, the sampling point value is also processed by filter patch, as shown in Fig. 2(b). The DLBP descriptor is formulated as:

$$DLBP_{P,R}(n_c) = \sum_{i=0}^{P-1} s(\Phi(N_{i,R,\omega_R}, \Phi(N_{i,R-1,\omega_{R-1}})) \cdot 2^i$$

$$(7)$$

where $N_{i,R-1,\omega_{R-1}}$ denotes the patch centered at the neighboring sampling point n_i with distance R-1 from the centered pixel n_c. Similar to the LBP, we also apply the strategy for obtaining $LBP_{P,R}^{riu2}$ to the DLBP. By computing DLBP, the relationship between interaction spaces in radial direction is highly considered. Similar to the LBP, it reflects macrostructural information when the radius is large. To fully extract both the microstructural and macrostructural information, in this paper, both the improved LBP and DLBP descriptors are
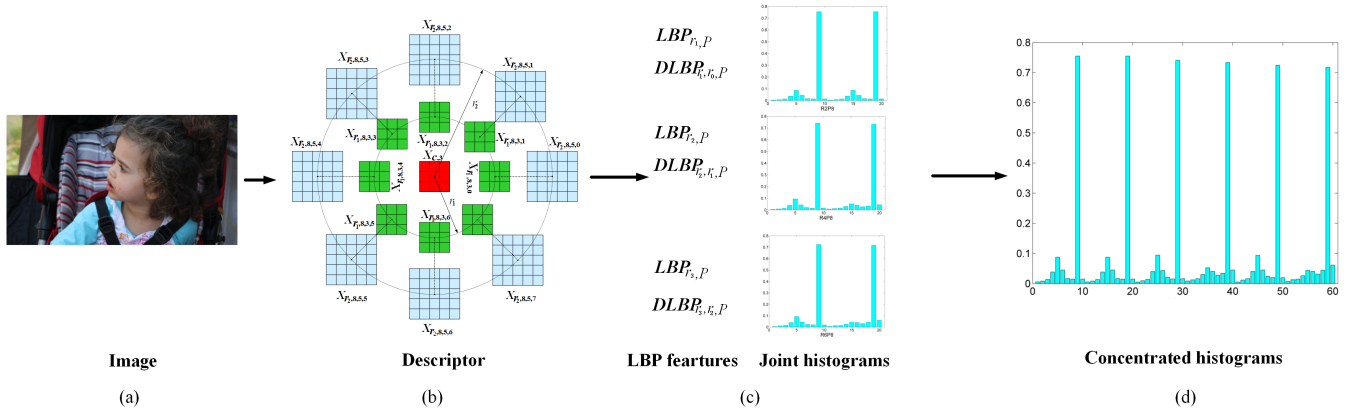
Fig. 2. Flow chart of feature extraction. (a) The input image; (b) The multi-scale improved LBP descriptor; (c) LBP and DLBP features at different scales and the corresponding joint histograms; (d) Final concentrated histograms.

operated with multiple radiuses. Fig. 2(b) illustrates the simple diagram of the proposed descriptor.

After the extraction of the LBP and DLBP feature maps, the frequency histogram of each map is calculated. Generally, a frequency histogram is composed of various elements and is defined as:

$$h_{j,M} = N_{j,M}/N_M, j = 1, 2, 3, \ldots, P+2 \quad (8)$$
$$M \in LBP, DLBP$$

where the subscript M denotes the category of the feature map. j is the selected value in the feature map with a maximum value of $P+2$ and a minimum value of 1. N_M is the pixel number of the corresponding feature map, while $N_{j,M}$ is the pixel number of the j-th value of feature map. Then, the joint histogram H_J is expressed by combining the LBP histogram and DLBP histogram together, $H_J = [H_{LBP}, H_{DLBP}]$. Finally, the concentrated histogram is combined by all joint histograms in multiple radiuses, $H_c = \{H_{J_1}, H_{J_2}, \cdots, H_{J_n}\}$, where H_{J_n} denotes the joint histogram when radius is n. Figs. 2(c)-(d) show the generation procedure of feature.

To manifest the potential of the proposed method on IQA problem, a simple test was conducted on an image named *baby* [1]. The image, as shown in Fig. 3, was first corrupted by Gaussian filter and then distorted by JPEG compression. The LBP feature maps of pristine and distorted images are depicted in Figs. 4(a)-(c). All the sub-figures keep to the following arrangement. The first row contains LBP maps (from the left to right represent maps with radius value of 2 to 6 at an interval of 2 pixels). The second row is the DLBP maps with the same arrangement as the first row. Through the comparison among these figures, we can find that the feature maps (the LBP and DLBP maps) change with respect to different distortions. To be more specific, the LBP and DLBP maps generated from the pristine image contain rich detailed information. While, the image structure is damaged and is full of block effects when it is mainly distorted by JPEG, as shown in Fig. 4(b). Likewise, the marginal zone is decayed when it is mainly distorted by Gblur, as shown in Fig. 4(c).

[1] The image is downloaded from Internet and the copyright belongs to its rightful owners. The authors do not claim ownership. No copyright infringement is intended.

To further demonstrate the distortion effect on joint histogram, we introduce another example, as shown in Fig. 5. We corrupt the *baby* image and obtain another two distorted images by Gblur+JPEG distortion. Fig. 5(a) represents the joint histograms of reference image (*baby*), while Figs. 5(b)-5(c) are the joint histograms of Gblur+JPEG distorted images. Each joint histogram contains two parts, which are the patterns of the LBP and DLBP. Among them, (1-10) represent the LBP patterns; (11-20) represent the DLBP patterns. Note that, compared to Fig. 5(b), Fig. 5(c) is the joint histograms of image that suffers from more serious Gblur distortion while contains equivalent intensity of JPEG distortion. As can be seen, with the distortion degree increased, the joint histogram changes. Specifically, taking the joint histograms of 'R2P8' as an example, patterns 1-4, 7, 11-14 and 17, gradually decline as the distortion degree increasing. Similar characteristics can be also observed in cases of 'R4P8' and 'R6P8'. Therefore, the joint histograms, to some extent, can reflect the degree of image distortion as does the concentrated histogram. Based on the above analysis, we have a reason to trust that the concentrated histograms can be used for solving the MD IQA problem.

Fig. 3. Three versions of the image named *baby*: (a) reference image, (b) and (c) are images corrupted by Gblur + JPEG. Note that (b) is mainly distorted by JPEG, while (c) is mainly distorted by Gblur.

B. Image Quality Estimation

Broadly speaking, mainstream NR-IQA metrics rely on the prior knowledge of the distortion category and degree. Machine learning is an effective technology to discover and exploit prior knowledge to build a good connection between the features extracted and the quality score to be predicted. Many NR-IQA metrics are proposed based on machine learning, and
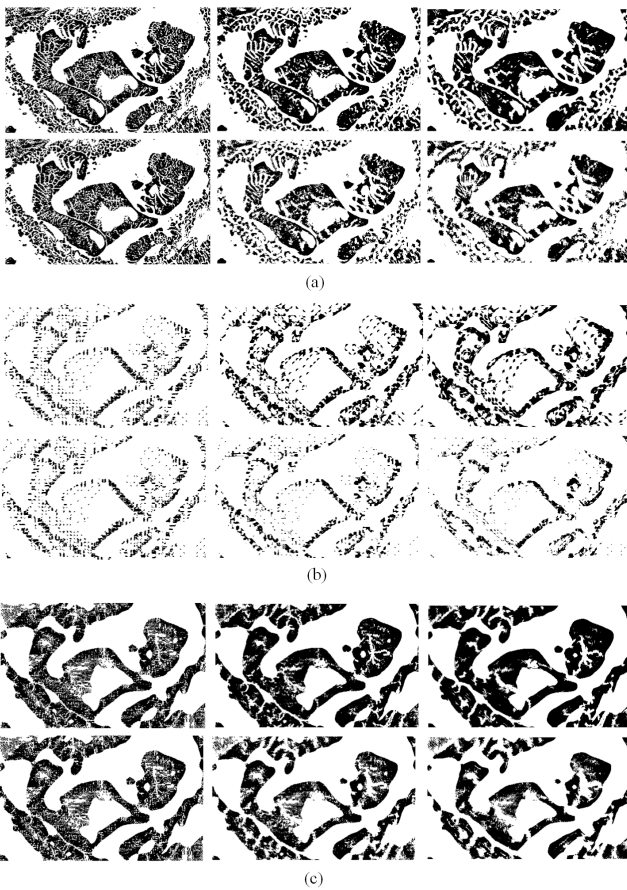
Fig. 4. Feature maps of Figs. 3. For each sub-figure, there are two rows. First row: LBP maps (from the left to right are maps with radius values of 2 to 6 at an interval of 2 pixels). Second row: DLBP maps (with the same arrangement rules as the first row).
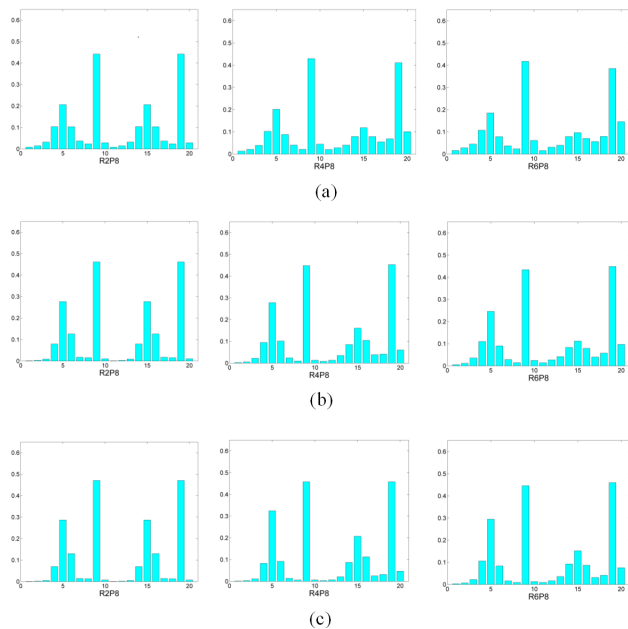

Fig. 5. Joint histograms (from the left to right are the histograms with radius values of 2 to 6 at an interval of 2 pixels). (a) Joint histograms of reference image (*baby*); (b)-(c) Joint histograms of Gblur + JPEG distorted image. Note that, compared to (b), (c) is calculated from image that suffers from more serious Gblur distortion while it contains the equivalent intensity to that of the JPEG distortion.

they usually operate in the subsequent three steps, including feature extraction, model generation and quality prediction, as shown in Fig. 6.
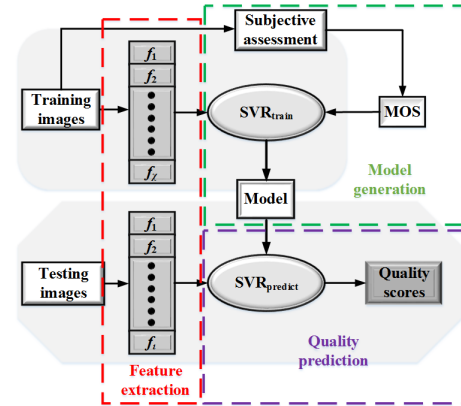

Fig. 6. The flowchart of the mainstream scheme of the learning-based NR-IQA metrics.

Among those three steps, the feature extraction is generally treated as the key point. For this point, in this paper, we improve the LBP descriptor by considering a greater radius to extract the microstructures and macrostructures of a given image signal. After feature extraction, a regression algorithm is used to learn a mapping from the feature vector (i.e., concentrated histogram H_c) to the quality score. Since the SVR has been widely proven to be efficient in the NR-IQA [10], [27], [33], this paper follows the previous works and chooses the libSVM package [34] to implement the SVR with the radial basis function as the kernel function. Notice that we extract 60 features and such high dimensionality of feature vector is easy to result in overfitting [35], [36]. Towards settling the overfitting problem, the random subspace method is introduced to be combined with the SVR for learning [36].

Then, given a test image, its quality can be predicted by feeding its extracted features into the trained model. After obtaining the predicted score, the performance of the proposed method is calculated across the predicted scores and the subjective ratings of test images. Since we do not require any information from the reference image, the proposed method belongs to the NR-IQA metric.

IV. EXPERIMENTAL RESULTS AND ANALYSIS

A. Experimental Protocol

1) Database Description: Two public MD image databases (MLIVE [37] and MDID 2013 [20]) are selected as the test platforms. The MLIVE database is composed of two parts, Part I and Part II. Each part contains 225 distorted images generated from 15 reference images. In detail, the images in Part I are corrupted by the Gblur+JPEG, while those in Part II are processed by the Gblur+WN. Besides, MLIVE also contains single distortion when the reference image is only corrupted by one distortion type. Different from the MLIVE database, the MDID 2013 database consists of images that are successively corrupted by three kinds of distortions (Gblur+JPEG+WN). A total of 324 distorted images are generated from 12 reference images.

2) Parameter Choice: There are three key parameters that affect the performance of the proposed method: the patch size, the radius value and the number of radii. In this paper, the patch is of size $\omega_R \times \omega_R$, where $\omega_R = R - 1$. Moreover, three-scale radiuses (R) are chosen with values from 2 to 6 at an interval of 2 pixels (more details are introduced in Section IV-D).

3) Performance Evaluation Criteria: Three commonly used criteria, i.e., Pearson linear correlation coefficient (PLCC), Spearman Rank order Correlation coefficient (SRCC) and root mean square error (RMSE) suggested by VQEG [38], were employed to evaluate and compare the proposed method with existing IQA metrics. To reduce the nonlinearity of the predicted score, a five-parameter logistic regression function is employed before PLCC and RMSE computation:

$$Q_p = \xi_1 \cdot \left(\frac{1}{2} - \frac{1}{exp(\xi_2 \cdot (q - \xi_3))}\right) + \xi_4 \cdot q + \xi_5 \quad (9)$$

where q is the raw score vector estimated using IQA metric; Q_p is the adopted quality vector; $\{\xi_1, \xi_2, \xi_3, \xi_4, \xi_5\}$ are the parameters of logistic function.

B. Performance Comparison with FR-IQA Methods

To investigate the effectiveness of the proposed method, a comparison was conducted with several state-of-the-art FR-IQA metrics, including PSNR, SSIM [39], ADD-SSIM [40], VSNR [41], Wang's metric [42], VIF [43], FSIM [44], IGM [45], and GMSD [46]. Since the proposed method is learning-based, we randomly divided the dataset into two parts (80% for training and 20% for testing) and report the median result of the performance across 100 iterations. Table I shows the experimental results. From Table I, we can find some meaningful observations. At first, some classical methods (e.g., PSNR, SSIM, and VSNR) generally obtain worse performance on the MDID 2013 database. Second, the proposed method achieves higher correlation and monotonicity with lower error on all the tested databases. Third, the performance of selected metrics on the MDID 2013 database is inferior to the MLIVE database. The reason for the above phenomena can be attributed to the following. First, conventional metrics are mainly designed for single distortion with the hypothesis that the extracted features have special relationship with the distortion. However, such a relationship may be broken when facing multiple distortions. Second, the image content becomes more complicated with less regularity as the increment of distorted category. Therefore, the performance on the MDID 2013 database is inferior to the performance on the MLIVE database reasonably.

C. Performance Comparison with NR-IQA Methods

In this subsection, we compare the proposed method with several mainstream NR-IQA metrics, including BRISQUE [9], BLIINDSII [12], NFERM [10], NIQE [11], IL-NIQE [47] and SSEQ [33]. For a fair comparison, all the learning-based metrics follow the rules with the same 80 % - 20 % database split and 100 iterations. The comparison results in the form of median values are listed in Table II.

From Table II, we can find that the proposed method performs consistently on the testing databases. It presents a higher predictive accuracy and monotonicity than other NR-IQA metrics. Similar to the comparison results in Section IV-B, the performance of the selected methods on the MLIVE is obviously superior to their counterparts on the MDID 2013. For example, the performances of some metrics (e.g., NFERM [10] and NIQE [11]) have relatively large difference. In addition, some selected NR-IQA metrics (e.g. NIQE [11] and IL-NIQE [47]) exhibit poor performance on the testing databases. In contrast, the proposed method obtains high performance on both databases with relatively small differences.

To further demonstrate the superiority of the proposed method, another comparison was carried out. The results between the proposed method and recently reported MD IQA metrics, including Li's [18], Lu's [17], GWH-GLBP [21], FISBLIM [48] and SISBLIM [20] are listed in Table III. All the results (except those of Li's [18] and Lu's [17]) are calculated using the released codes provided by authors. For comparison, both Li's and Lu's performance are directly taken from the associated papers. Since only three bits after decimal point are retained in Li's and Lu's works, we also processed other metrics' results accordingly. As seen, the proposed method has obtained very encouraging results, outperforming all the selected MD IQA metrics in terms of both monotonicity and accuracy on the MDID 2013, MLIVE I and MLIVE II databases. In particular, on the MLIVE I database, the SRCC of proposed method exceeds 0.038 more than the GWH-GLBP, which has obtained the best performance on the MD IQA problem so far. Similar results can be obtained with improvement of 0.013 on the MDID 2013 database. Although the improvement is smaller, it is of vital importance for the high-performance IQA. However, the proposed method is slightly inferior to GWH-GLBP with respect to the PLCC and RMSE on the whole MLIVE database. This may be because it increases the difficulty of distinguishing the degradations caused by Gblur+JPEG and Gblur+WN when MLIVE I and MLIVE II are combined together. Since the GWH-GLBP is operated on the gradient domain and utilizes the gradient magnitude as the weight, it can well distinguish the degradations by putting more emphasis on the image's structural information. In spite of this, its performance is inferior to the proposed method when the degradation becomes more complex (e.g., with three distortion types). In summary, the proposed method has a good ability to evaluate the quality of MD images.

TABLE III
SUMMARY OF EXPERIMENTAL RESULTS OF THE PROPOSED METHOD AND COMPETING MD IQA METRICS

Database	Metrics	Li's	Lu's	GWH-GLBP	FISBLIM	SISBLIM	Pro. Method
MDID 2013	SRCC	-	-	0.931	0.772	0.808	**0.944**
	PLCC	-	-	0.941	0.734	0.811	**0.948**
	RMSE	-	-	0.017	0.035	0.030	**0.016**
MLIVE I	SRCC	0.887	0.908	0.920	0.853	0.868	**0.958**
	PLCC	0.923	0.948	0.954	0.884	0.894	**0.962**
	RMSE	7.300	-	5.773	8.813	8.462	**4.424**
MLIVE II	SRCC	0.888	0.904	0.935	0.864	0.891	**0.948**
	PLCC	0.900	0.937	0.942	0.883	0.902	**0.957**
	RMSE	7.921	-	6.443	8.911	8.182	**4.963**
MLIVE	SRCC	0.885	-	**0.954**	0.856	0.878	0.954
	PLCC	0.901	-	**0.960**	0.880	0.895	0.959
	RMSE	8.000	-	**5.121**	8.962	10.291	5.128

To draw a more reliable conclusion, the statistical analysis

TABLE I
PERFORMANCE COMPARISONS WITH STATE-OF-THE-ART FR-IQA METRICS

Database	Metrics	PSNR	SSIM	ADD-SSIM	VSNR	Wang's	VIF	FSIM	IGM	GMSD	Pro. Method
MDID 2013	SRCC	0.5681	0.4890	0.7482	0.3906	0.8159	0.8301	0.6436	0.8095	0.8291	**0.9439**
	PLCC	0.5604	0.4873	0.7984	0.4335	0.7790	0.8444	0.5817	0.8232	0.8283	**0.9475**
	RMSE	0.0418	0.0443	0.0337	0.0468	0.0315	0.0283	0.0389	0.0298	0.0284	**0.0162**
MLIVE I	SRCC	0.7392	0.8970	0.8892	0.8366	0.8805	0.9196	0.8322	0.8933	0.8766	**0.9578**
	PLCC	0.7015	0.8494	0.8792	0.7877	0.8924	0.8788	0.8544	0.8559	0.8497	**0.9621**
	RMSE	12.9036	8.4694	8.7678	10.4978	8.4394	7.5254	10.6238	8.6134	9.2218	**4.4237**
MLIVE II	SRCC	0.7764	0.8947	0.8844	0.8052	0.8669	0.8945	0.8805	0.8831	0.8690	**0.9482**
	PLCC	0.7088	0.8760	0.8583	0.7652	0.8767	0.8807	0.8642	0.8548	0.8420	**0.9571**
	RMSE	11.7558	8.3311	8.7067	11.0628	8.8939	8.3400	8.8416	8.7522	9.2301	**4.9632**
MLIVE	SRCC	0.7456	0.8915	0.8838	0.8117	0.8736	0.9030	0.8933	0.8859	0.8817	**0.9541**
	PLCC	0.6965	0.8604	0.8566	0.7727	0.8852	0.8823	0.8635	0.8562	0.8519	**0.9592**
	RMSE	12.6021	8.5692	8.8489	11.0454	8.4451	8.6582	8.4990	8.7744	8.9221	**5.1282**

TABLE II
PERFORMANCE COMPARISONS WITH STATE-OF-THE-ART NR-IQA METRICS

Database	Metrics	BRISQUE	BLIINDS-II	NFERM	NIQE	IL-NIQE	SSEQ	Pro. Method
MDID 2013	SRCC	0.8734	0.9145	0.8902	0.5704	0.5156	0.9100	**0.9439**
	PLCC	0.8901	0.9218	0.9030	0.5724	0.5149	0.9194	**0.9475**
	RMSE	0.0227	0.0193	0.0219	0.0417	0.0435	0.0194	**0.0162**
MLIVE I	SRCC	0.9110	0.8955	0.9217	0.9102	0.9049	0.8983	**0.9578**
	PLCC	0.9405	0.9307	0.9488	0.8724	0.8914	0.9310	**0.9621**
	RMSE	6.3934	6.8729	5.9901	7.9338	8.1565	7.0283	**4.4237**
MLIVE II	SRCC	0.8696	0.8867	0.8890	0.8482	0.8970	0.9077	**0.9482**
	PLCC	0.9046	0.9137	0.9160	0.7944	0.8825	0.9235	**0.9571**
	RMSE	7.7549	7.4447	7.3760	9.8810	8.2472	7.1430	**4.9632**
MLIVE	SRCC	0.9022	0.9089	0.9170	0.8391	0.8912	0.9059	**0.9541**
	PLCC	0.9260	0.9237	0.9386	0.7750	0.8777	0.9212	**0.9592**
	RMSE	7.0193	7.2432	6.5263	10.2884	8.5778	7.2639	**5.1282**

was conducted for the comparison among various NR-IQA metrics. Since both the NIQE [11] and IL-NIQE [47] are free of learning, only five metrics are selected for further validation. The analysis of variance test was conducted to validate whether two IQA metrics had significant difference. Table IV details the test results. For simplicity, we respectively use number 1 to 6 to represent the proposed method, NFERM [10], GWH-GLBP [21], BRISQUE [9], BLIINDS-II [12] and SSEQ [33]. The symbol values of '1', '−' and '0' indicate that the method in the row is significantly better, comparable or worse than the compared metric in that column (with 95% confidence), respectively. In Table IV, each entry corresponds to the results of the statistical test performed on the criteria (from left to right: PLCC, SRCC, and RMSE). For instance, the second 1 in entry "111" in line 1, column 2 indicates the significant effect on SRCC. From Table I to Table IV, we can conclude that the proposed method outperforms the existing metrics in most cases. 1) It has the best performance on all the selected databases. 2) The top three methods on all the databases are the proposed method and GWH-GLBP [21], and NFERM [10]. Specifically, NFERM [10] achieves a mediocre performance in MDID 2013 database and a better performance on the MLIVE database. Therefore, it is not robust. In contrast, the proposed method exhibits relatively high and stable performances on all the databases. Moreover, it outperforms the GWH-GLBP [21] on the MLIVE II database. 3) In summary, our metric was significantly superior to other metrics 49 times, and comparable with them 11 times in the 60 comparisons. To be specific, our method outperforms all other metrics in MLIVE I, while it contains the advantage over most of metrics in other databases. These findings supplement Tables I and II in order to draw the conclusion that the proposed method performs consistently well in all MD image databases and outperforms most IQA metrics.

TABLE IV
STATISTICAL ANALYSIS AMONG DIFFERENT METRICS

	MDID2013						MLIVE I					
Metric	1	2	3	4	5	6	1	2	3	4	5	6
1	—	111	—	111	111	111	—	—	1–	111	111	111
2	000	—	000	000	000	000	000	—	—	111	111	111
3	000	111	—	111	111	111	0–	—	—	111	111	111
4	000	111	000	—	000	—	000	000	000	—	—	—
5	000	111	000	—	—	—	000	000	000	—	—	—
6	000	111	000	—	—	—	000	000	000	—	—	—
	MLIVE II						MLIVE					
Metric	1	2	3	4	5	6	1	2	3	4	5	6
1	—	111	—	111	111	111	—	111	—	111	111	111
2	000	—	000	0–	—	0–	000	—	000	111	111	111
3	—	111	—	111	111	111	—	111	—	111	111	111
4	000	1–	000	—	0–	—	000	000	000	—	—	—
5	000	—	000	0–	—	0–	000	000	000	—	—	—
6	000	1–	000	—	1–	—	000	1–	000	—	—	—

D. Variation with Algorithm Parameters

Since visual acuity declines in an orderly fashion with eccentricity [49], the radius value in the proposed method cannot be infinitely large. As discussed in Section IV-A, the number of radii, the radius value and patch size of the proposed method have direct impacts on the final performance. Therefore, it is meaningful to explore the relationship between these characteristics and the performance. Because the spatial interaction zone appears to be an elliptical shape [30], for simplicity, we apply the square patch to represent it. Since it is radially elongated, the patch size increases as the radius (R)

increases. In this paper, the patch size is empirically set as $\omega_R \times \omega_R$, where $\omega_R = R - 1$. Meanwhile, the minimal radius value is set as 2 pixels and increases at an interval of 2 pixels. Thus, the radius number is the only factor that affects the final results. Table V tabulates the PLCC value on MD image databases under different radius numbers. Some meaningful observations can be derived from the table as follows. 1) The proposed method obtains a mediocre performance when the radius number is 1. The reason can be attributed to the fact that, only microstructure information is included when the radius is small. 2) The PLCC value increases as the radius number increases. Specifically, the proposed method achieves encouraging results on all databases, even though the number is only 2. In particular, the PLCC value exceeds 0.950 on the MLIVE I database and outperforms most metrics. 3) After reaching the peak (when the radius number is 3), the PLCC value then slowly declines. In summary, the performance is always competitive with respect to the other metrics when the radius number changes from 2 to 4, which demonstrates the robustness of the proposed method. Therefore, in this paper, the radius number is set as 3.

TABLE V
NUMBERS OF RADIUS EFFECT ON PLCC

Database	Number of Radius Number			
	1	2	3	4
MDID 2013	0.9195	0.9448	0.9475	0.9159
MLIVE I	0.9332	0.9502	0.9621	0.9483
MLIVE II	0.8972	0.9382	0.9571	0.9211
MLIVE	0.9052	0.9291	0.9592	0.9257

E. Performance on Authentic Image Distortions

As previously proven, the proposed method achieves encouraging performance on both the MLIVE and MDID databases. However, both of them are generated by using synthetic distortions through combining different types of human-made distortions. Advanced image capture devices, including mobile phones, digital video cameras and wearable devices, allow us frequently and conveniently to acquire images. Therefore, it is meaningful and interesting to design an objective assessment method for distorted images by commercial consumer products [50]. In this paper, we further test and compare the proposed method with several NR-IQA methods on the CLIVE (LIVE in the wild image quality challenge database) database [51], [52]. It includes 1162 distorted images captured using typical real-world mobile cameras with various authentic image distortions, such as low-light blur, noise, motion blur, overexposure, underexposure compression errors and their various combinations. Table VI shows the experimental results. As seen, all traditional metrics merely obtain a general performance (e.g., PLCC is less than 0.63), while the specific metric (FRIQUEE [4]) for authentic image distortions obtains the best performance. The mediocre performance indicates that traditional IQA metrics are far from the usefulness in the real field. Fortunately, the proposed method is superior to all competing traditional metrics, which strengthens our confidence and plays a guiding role in future work. Although

difficulties exist in tackling the quality assessment problem of authentic image distortions, as the ultimate aim, it is the challenge of the infinite charm of the IQA that attracts us to conquer it.

F. Discussions

The MD IQA is faced with great challenges. Inspired by the human visual mechanism, we propose an effective NR-IQA metric. Our method has been shown to achieve encouraging performance on the MLIVE and MDID2013 databases. The superior performance originates from using the IMLBP to extract both the local structural and macrostructural information. Although the proposed method has achieved a fairly good performance on the MD IQA, it merely obtains general performance on authentic distortions as with other IQA metrics. The performance can be further improved through the following attempts. First, other LBP descriptors (such as the neighborhood intensity LBP) can be integrated into the proposed method and form more robustness descriptor. Second, the patch filter can be replaced with other efficient filters, for instance the Gaussian filter, the maximum filter and the mean filter, to further improve the results. Third, more features that represent image details, such as global structural information, color information, frequency information and more, can be introduced to achieve better performance results. In the future, we plan to incorporate the features used in our model and those extracted from FRIQUEE to propose a higher-performance blind quality metric.

V. CONCLUSION

In this paper, we present an effective NR-IQA algorithm to estimate the MD image quality. Inspired by the human visual mechanism, the improved multi-scale LBP descriptor is proposed to extract the local microstructural and macrostructural information. Then the statistical histograms of the feature maps that are highly relevant to the image quality are calculated as the input of the SVR model for training the effective IQA metric. The effectiveness is validated on public MD image databases. The experimental results demonstrate that the proposed method outperforms most selected state-of-the-art FR-IQA and NR-IQA metrics. Moreover, the results on the CLIVE database further validate its potential for the quality assessment of authentic image distortions.

REFERENCES

[1] H. R. Wu, A. R. Reibman, W. Lin, F. Pereira, and S. S. Hemami, "Perceptual visual signal compression and transmission," *Proceedings of the IEEE*, vol. 101, no. 9, pp. 2025–2043, 2013.

[2] W. Lin and C. C. J. Kuo, "Perceptual visual quality metrics: A survey," *Journal of Visual Communication & Image Representation*, vol. 22, no. 4, pp. 297–312, 2011.

[3] A. C. Bovik, "Automatic prediction of perceptual image and video quality," *Proceedings of the IEEE*, vol. 101, no. 9, pp. 2008–2024, 2013.

[4] D. Ghadiyaram and A. C. Bovik, "Perceptual quality prediction on authentically distorted images using a bag of features approach," *Journal of Vision*, vol. 17, no. 1, pp. 1–25, 2017.

[5] K. Gu, J. Qiao, X. Min, G. Yue, W. Lin, and D. Thalmann, "Evaluating quality of screen content images via structural variation analysis," *IEEE Transactions on Visualization and Computer Graphics*, 2017, DOI: 10.1109/TVCG.2017.2771284.

TABLE VI
PERFORMANCE COMPARISONS ON CLIVE DATABASE WITH STATE-OF-THE-ART NR-IQA METRICS

Database	Metrics	BRISQUE	BLIINDS-II	NFERM	GWH-GLBP	NIQE	IL-NIQE	SSEQ	FISBLIM	SISBLIM	FRIQUEE	Pro. Method
CLIVE	SRCC	0.5872	0.4036	0.5403	0.6105	0.4232	0.4393	0.4269	0.4050	0.4289	0.7201	0.6207
	PLCC	0.6154	0.4547	0.5705	0.6118	0.4834	0.5047	0.4523	0.4723	0.4470	0.7203	0.6289

[6] G. Yue, C. Hou, K. Gu, and N. Ling, "No reference image blurriness assessment with local binary patterns," *Journal of Visual Communication and Image Representation*, vol. 49, pp. 382–391, 2017.

[7] Q. Jiang, F. Shao, W. Lin, K. Gu, G. Jiang, and H. Sun, "Optimizing multi-stage discriminative dictionaries for blind image quality assessment," *IEEE Transactions on Multimedia*, 2017, DOI: 10.1109/TM-M.2017.2763321.

[8] G. Yue, C. Hou, K. Gu, S. Mao, and W. Zhang, "Biologically inspired blind quality assessment of tone-mapped images," *IEEE Transactions on Industrial Electronics*, vol. 65, no. 3, pp. 2525–2536, 2018.

[9] A. Mittal, A. K. Moorthy, and A. C. Bovik, "No-reference image quality assessment in the spatial domain." *IEEE Transactions on Image Processing*, vol. 21, no. 12, pp. 4695–4708, 2012.

[10] K. Gu, G. Zhai, X. Yang, and W. Zhang, "Using free energy principle for blind image quality assessment," *IEEE Transactions on Multimedia*, vol. 17, no. 1, pp. 50–63, 2015.

[11] A. Mittal, R. Soundararajan, and A. C. Bovik, "Making a "completely blind" image quality analyzer," *IEEE Signal Processing Letters*, vol. 20, no. 3, pp. 209–212, 2013.

[12] M. A. Saad, A. C. Bovik, and C. Charrier, "Blind image quality assessment: A natural scene statistics approach in the DCT domain," *IEEE Transactions on Image Processing*, vol. 21, no. 8, pp. 3339–52, 2012.

[13] W. Lu, K. Zeng, D. Tao, Y. Yuan, and X. Gao, "No-reference image quality assessment in contourlet domain," *Neurocomputing*, vol. 73, no. 4, pp. 784–794, 2010.

[14] A. K. Moorthy and A. C. Bovik, "A two-step framework for constructing blind image quality indices," *IEEE Signal Processing Letters*, vol. 17, no. 5, pp. 513–516, 2010.

[15] P. Bouboulis, S. Theodoridis, C. Mavroforakis, and L. Evaggelatoudalla, "Complex support vector machines for regression and quaternary classification." *IEEE Transactions on Neural Networks & Learning Systems*, vol. 26, no. 6, pp. 1260–1274, 2014.

[16] S. C. Pei and L. H. Chen, "Image quality assessment using human visual DOG model fused with random forest," *IEEE Transactions on Image Processing*, vol. 24, no. 11, pp. 3282–92, 2015.

[17] Y. Lu, F. Xie, T. Liu, and Z. Jiang, "No reference quality assessment for multiply-distorted images based on an improved bag-of-words model," *IEEE Signal Processing Letters*, vol. 22, no. 10, pp. 1811–1815, 2015.

[18] C. Li, Y. Zhang, X. Wu, W. Fang, and L. Mao, "Blind multiply distorted image quality assessment using relevant perceptual features," in *IEEE International Conference on Image Processing*, 2015, pp. 4883–4886.

[19] D. M. Chandler, "Seven challenges in image quality assessment: past, present, and future research," *ISRN Signal Processing*, vol. 2013, 2013.

[20] K. Gu, G. Zhai, X. Yang, and W. Zhang, "Hybrid no-reference quality metric for singly and multiply distorted images," *IEEE Transactions on Broadcasting*, vol. 60, no. 3, pp. 555–567, 2014.

[21] Q. Li, W. Lin, and Y. Fang, "No-reference quality assessment for multiply-distorted images in gradient domain," *IEEE Signal Processing Letters*, vol. 23, no. 4, pp. 541–545, 2016.

[22] T. Ojala, M. Pietikainen, and T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Transactions on Pattern Analysis & Machine Intelligence*, vol. 24, no. 7, pp. 971–987, 2002.

[23] B. Zhang, Y. Gao, S. Zhao, and J. Liu, "Local derivative pattern versus local binary pattern: face recognition with high-order local pattern descriptor," *IEEE Transactions on Image Processing*, vol. 19, no. 2, pp. 533–544, 2010.

[24] Z. Guo, X. Wang, J. Zhou, and J. You, "Robust texture image representation by scale selective local binary patterns," *IEEE Transactions on Image Processing*, vol. 25, no. 2, pp. 687–699, 2015.

[25] Q. Wu, H. Li, F. Meng, and K. N. Ngan, "Blind image quality assessment based on multichannel feature fusion and label transfer," *IEEE Transactions on Circuits & Systems for Video Technology*, vol. 26, no. 3, pp. 425–440, 2016.

[26] M. Zhang, C. Muramatsu, X. Zhou, T. Hara, and H. Fujita, "Blind image quality assessment using the joint statistics of generalized local binary

pattern," *IEEE Signal Processing Letters*, vol. 22, no. 2, pp. 207–210, 2015.

[27] S. Du, Y. Yan, and Y. Ma, "Blind image quality assessment with the histogram sequences of high-order local derivative patterns," *Digital Signal Processing*, vol. 55, pp. 1–12, 2016.

[28] D. G. Pelli and K. A. Tillman, "The uncrowded window of object recognition," *Nature Neuroscience*, vol. 11, no. 10, pp. 1129–1135, 2008.

[29] J. Intriligator and P. Cavanagh, "The spatial resolution of visual attention," *Cognitive Psychology*, vol. 43, no. 3, pp. 171–216, 2001.

[30] A. Toet and D. M. Levi, "The two-dimensional shape of spatial interaction zones in the parafovea," *Vision Research*, vol. 32, no. 7, pp. 1349–1357, 1992.

[31] B. Guenter, M. Finch, S. Drucker, D. Tan, and J. Snyder, "Foveated 3D graphics," *ACM Transactions on Graphics*, vol. 31, no. 6, p. 164, 2012.

[32] S. Lee, M. S. Pattichis, and A. C. Bovik, "Foveated video compression with optimal rate control," *IEEE Transactions on Image Processing*, vol. 10, no. 7, pp. 977–992, 2001.

[33] L. Liu, B. Liu, H. Huang, and A. C. Bovik, "No-reference image quality assessment based on spatial and spectral entropies," *Signal Processing Image Communication*, vol. 29, no. 8, pp. 856–863, 2014.

[34] C.-C. Chang and C.-J. Lin, "LIBSVM: a library for support vector machines," *ACM Transactions on Intelligent Systems and Technology*, vol. 2, no. 3, pp. 1–27, 2011.

[35] T. K. Ho, "The random subspace method for constructing decision forests," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, no. 8, pp. 832–844, 1998.

[36] D. Tao, X. Tang, X. Li, and X. Wu, "Asymmetric bagging and random subspace for support vector machines-based relevance feedback in image retrieval," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 7, pp. 1088–1099, 2006.

[37] D. Jayaraman, A. Mittal, A. K. Moorthy, and A. C. Bovik, "Objective quality assessment of multiply distorted images," in *Conference Record of the Forty Sixth Asilomar Conference on Signals, Systems and Computers*. IEEE, 2012, pp. 1693–1697.

[38] T. V. Q. E. Group, "Final report from the video quality experts group on the validation of objective models of video quality assessment, phase ii (fr-tv2)," *ftp://ftp.its.bldrdoc.gov/dist/ituvidq/Boulder-VQEG-jan-04/VQEG-Phase II-FRTV-Final-Report-SG9060E.doc*.

[39] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600 – 612, 2004.

[40] K. Gu, S. Wang, G. Zhai, and W. Lin, "Analysis of distortion distribution for pooling in image quality prediction," *IEEE Transactions on Broadcasting*, vol. 62, no. 2, pp. 446–456, 2016.

[41] D. M. Chandler and S. S. Hemami, "VSNR: A wavelet-based visual signal-to-noise ratio for natural images," *IEEE transactions on image processing*, vol. 16, no. 9, pp. 2284–2298, 2007.

[42] S. Wang, C. Deng, W. Lin, G. B. Huang, and B. Zhao, "NMF-based image quality assessment using extreme learning machine," *IEEE Transactions on Cybernetics*, vol. 47, no. 1, pp. 232–243, 2017.

[43] H. R. Sheikh and A. C. Bovik, "Image information and visual quality," *IEEE Transactions on Image Processing*, vol. 15, no. 2, pp. 430–444, 2006.

[44] L. Zhang, L. Zhang, X. Mou, and D. Zhang, "FSIM: a feature similarity index for image quality assessment." *IEEE Transactions on Image Processing*, vol. 20, no. 8, pp. 2378–2386, 2011.

[45] J. Wu, W. Lin, G. Shi, and A. Liu, "Perceptual quality metric with internal generative mechanism," *IEEE Transactions on Image Processing*, vol. 22, no. 1, pp. 43–54, 2013.

[46] W. Xue, L. Zhang, X. Mou, and A. C. Bovik, "Gradient magnitude similarity deviation: A highly efficient perceptual image quality index," *IEEE Transactions on Image Processing*, vol. 23, no. 2, pp. 684–95, 2014.

[47] L. Zhang, L. Zhang, and A. C. Bovik, "A feature-enriched completely blind image quality evaluator." *IEEE Transactions on Image Processing*, vol. 24, no. 8, pp. 2579–2591, 2015.

[48] K. Gu, G. Zhai, M. Liu, X. Yang, W. Zhang, X. Sun, W. Chen, and Y. Zuo, "FISBLIM: A five-step blind metric for quality assessment of multiply distorted images." in *SiPS*, 2013, pp. 241–246.

[49] P. J. Bennett and M. S. Banks, "Sensitivity loss in odd-symmetric mechanisms and phase anomalies in peripheral vision." *Nature*, vol. 326, no. 6116, pp. 873–6, 1987.

[50] D. Ghadiyaram and A. C. Bovik, "Scene statistics of authentically distorted images in perceptually relevant color spaces for blind image quality assessment," in *IEEE International Conference on Image Processing*. IEEE, 2015, pp. 3851–3855.

[51] ——, "Massive online crowdsourced study of subjective and objective picture quality," *IEEE Transactions on Image Processing*, vol. 25, no. 1, pp. 372–387, 2016.

[52] ——, "Massive online crowdsourced study of subjective and objective picture quality," *Online: http://live.ece.utexas.edu/research/ChallengeDB/index.html*, 2015.

Nam Ling Nam Ling (S'85-M'88-SM'99-F'08) received the B.Eng. degree from the National University of Singapore, Singapore, in 1981, and the M.S. and Ph.D. degrees from the University of Louisiana, Lafayette, LA, USA, in 1985 and 1989, respectively.

From 2002 to 2010, he was an Associate Dean with the School of Engineering, Santa Clara University, Santa Clara, CA, USA. He is currently the Sanfilippo Family Chair Professor and the Chair of the Department of Computer Engineering, Santa Clara University. He is also a Consulting Professor with the National University of Singapore; a Guest Professor with Tianjin University, Tianjin, China; a Guest Professor with Shanghai Jiao Tong University, Shanghai, China; a Cuiying Chair Professor with Lanzhou University, Gansu, China; and a Distinguished Professor with the Xian University of Posts and Telecommunications, Xian, China. He has authored or coauthored more than 190 publications and standard contributions, including two books in the fields of video coding and systolic arrays. He has filed/granted more than 20 U.S. patents.

Dr. Ling is an IET Fellow. He has served as a General Chair/Co-Chair of IEEE Hot Chips, VCVP (twice), IEEE ICME, IEEE Umedia (five times), and IEEE SiPS. He has also served as a Technical Program Co-Chair of IEEE ISCAS, APSIPA ASC, IEEE APCCAS, IEEE SiPS (twice), DCV, and IEEE VCIP. He was a Technical Committee Chair of IEEE CASCOM TC and IEEE TCMM, and has served as a Guest Editor or Associate Editor of the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMSH: REGULAR PAPERS, the IEEE JOURNAL OF SELECTED TOPICS IN SIGNAL PROCESSING, Springer JSPS, Springer MSSP, and more. He was named an IEEE Distinguished Lecturer twice and was also an APSIPA Distinguished Lecturer. He was the recipient of the IEEE ICCE Best Paper Award (first place) and the IEEE Umedia Best Paper Award. He was the recipient of six awards from Santa Clara University, four at the university level (Outstanding Achievement, Recent Achievement in Scholarship, Presidents Recognition, and Sustained Excellence in Scholarship) and two at the school/college level (Researcher of the Year and Teaching Excellence). He was a Keynote Speaker of IEEE APCCAS, VCVP (twice), JCPC, IEEE ICAST, IEEE ICIEA, IET FC Umedia, IEEE Umedia, IEEE ICCIT, and Workshop at XUPT (twice), as well as a Distinguished Speaker of IEEE ICIEA.
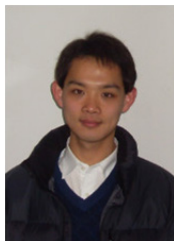
Guanghui Yue received the B.S. degree in communication engineering from Tianjin University, Tianjin, China, in 2014, and he is currently working toward the Ph.D. degree in the School of Electrical and Information Engineering, Tianjin University, China. He also has been a joint Ph.D. student with School of Computer Science and Engineering, Nanyang Technological University since September, 2017.

His research interests include bioelectrical signal processing, multimedia quality assessment, 3-D image visual discomfort prediction and pattern recognition.

Chunping Hou received the M.Eng. and Ph.D. degrees in electronic engineering from Tianjin University, Tianjin, China, in 1986 and 1998, respectively. Since 1986, she has been with the School of Electronic and Information Engineering, Tianjin University, Tianjin, China, where she is currently a Full Professor and the Director of the Broadband Wireless Communications and 3-D Imaging Institute.

Her current research interests include 3-D image processing, 3-D display, wireless communication, and the design and applications of communication systems.

Beichen Li received B.S.degree from University of Electronic Science and Technology, Chengdu, China, then got his M.S. degree from Steven Institute of Technology at Hoboken, U.S.. He has worked as a software developer and researcher in multiple organizations include Good Tech, NeuSoft, KingSoft Cloud, Bright Oceans Inter-Telecom Coporation and Maritime Security Lab. Beichen's research interests include cloud computing, big data, computer vision and machine learning.

Ke Gu received the B.S. and PhD degrees in electronic engineering from Shanghai Jiao Tong University, Shanghai, China, in 2009 and 2015. He is the reviewer for IEEE T-IP, T-MM, T-CSVT, T-CYB, T-BC, J-STSP, SPL, Information Sciences, Neurocomputing, SPIC, JVCI, DSP, etc. He has reviewed more than 30 journal papers each year. His research interests include quality assessment, contrast enhancement and visual saliency detection. Dr. Gu received the Best Paper Award at the IEEE International Conference on Multimedia and Expo (ICME) 2016. He is a Special Session Organizer for VCIP 2016.