No-Reference and Robust Image Sharpness Evaluation Based on Multiscale Spatial and Spectral Features

Leida Li, Wenhan Xia, Weisi Lin, *Fellow, IEEE*, Yuming Fang, and Shiqi Wang

***Abstract*—The human visual system exhibits multiscale characteristic when perceiving visual scenes. The hierarchical structures of an image are contained in its scale space representation, in which the image can be portrayed by a series of increasingly smoothed images. Inspired by this, this paper presents a no-reference and robust image sharpness evaluation (RISE) method by learning multiscale features extracted in both the spatial and spectral domains. For an image, the scale space is first built. Then sharpness-aware features are extracted in gradient domain and singular value decomposition domain, respectively. In order to take into account the impact of viewing distance on image quality, the input image is also down-sampled by several times, and the DCT-domain entropies are calculated as quality features. Finally, all features are utilized to learn a support vector regression model for sharpness prediction. Extensive experiments are conducted on four synthetically and two real blurred image databases. The experimental results demonstrate that the proposed RISE metric is superior to the relevant state-of-the-art methods for evaluating both synthetic and real blurring. Furthermore, the proposed metric is robust, which means that it has very good generalization ability.**

***Index Terms*—Image sharpness evaluation, scale space, gradient, singular value decomposition, entropy, support vector regression (SVR).**

I. Introduction

VISUAL quality evaluation has been an increasingly important problem in multimedia processing and communication systems [1]–[3]. Since human is the end consumer of multimedia signals, the most straightforward way to quality evaluation is subjective rating. However, subjective evaluation is laborious and cannot be integrated into real-time applications. In the past several years, objective quality evaluation has drawn extensive attention, which is to build computational models for measuring image degradations and meanwhile maintain consistency with the human perception [4], [5]. Quality assessment is becoming increasingly important for many image processing applications, including compression [6]–[8], enhancement [9], [10], restoration [11] and image forensics [12], [13], etc. According to the availability of reference information, the existing quality metrics can be categorized into full-reference (FR) [14], reduced-reference (RR) [15], [16] and no-reference (NR) [17], [18] approaches.

Sharpness/blur is one of the most important factors of image quality, which is very likely to occur across the whole lifecycle of images, e.g., acquisition, processing, storage and transmission. Objective image sharpness evaluation is crucial for modern imaging devices, and it also plays a vital role in benchmarking image processing algorithms. In real-world applications, reference images are usually not available, so sharpness evaluation needs to be done in a NR manner.

In the literature, several NR image sharpness/blur metrics have been proposed. Marziliano *et al.* [19] proposed a simple method by measuring the spread of image edges in the spatial domain. Edges were first detected using the Sobel edge detector. Then the sharpness index was calculated as the mean edge width across the whole image. In [20], Ferzli *et al.* proposed the Just Noticeable Blur (JNB) model. JNB was first determined based on image local contrast. Then it was combined into a probability summation model for generating the final blur score. Narvekar *et al.* [21] addressed a method to estimate the probability of detecting blur using a probabilistic model. Then the information was pooled by calculating the cumulative probability of blur detection (CPBD), producing the sharpness score. Vu *et al.* [22] proposed the spectral and spatial sharpness (S3) model. For an image, the magnitude spectrum slop was utilized to estimate the attenuation of high-frequency information; while the Total Variation (TV) model was adopted to compensate for the impact of contrast component on sharpness perception. Based on these features, a sharpness map was generated and the final sharpness score was generated using percentile pooling. In [23], the authors proposed a fast image sharpness (FISH) model based on Discrete Wavelet Transform (DWT). Given an image, it was first transformed into the DWT domain. Then the log-energies of the DWT subbands were computed

Manuscript received September 3, 2016; revised November 23, 2016; accepted December 10, 2016. Date of publication December 15, 2016; date of current version April 15, 2017. This work was supported in part by the National Natural Science Foundation of China under Grant 61379143, in part by the Fundamental Research Funds for the Central Universities under Grant 2015XKMS032, and in part by the Qing Lan Project. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Lingfen Sun. *(Corresponding author: Yuming Fang.)*

L. Li and W. Xia are with the School of Information and Electrical Engineering, China University of Mining and Technology, Xuzhou 221116, China (e-mail: reader1104@hotmail.com; 969872673@qq.com).

W. Lin and S. Wang are with the School of Computer Science and Engineering, Nanyang Technological University, Singapore 639798 (e-mail: wslin@ntu.edu.sg; sqwang1986@gmail.com).

Y. Fang is with the School of Information Technology, Jiangxi University of Finance and Economics, Nanchang 330032, China (e-mail: fa0001ng@e.ntu.edu.sg).

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

Digital Object Identifier 10.1109/TMM.2016.2640762

and their weighted average value was defined as the overall sharpness score. Sang *et al.* [24] addressed a simple method based on the singular value curve (SVC). Hassen *et al.* [25] proposed a sharpness index using the local phase coherence (LPC-SI), which is sensitive to image sharpness changes. In [26], Bahrami *et al.* proposed to use the maximum local variation (MLV), which was defined as the maximum gray-scale variation of a pixel in a small neighborhood. Gu *et al.* [27] proposed an autoregressive (AR) based image sharpness metric (ARISM), which was achieved via the analysis of the AR parameter space of images. In [28], the authors proposed a simple and effective image sharpness model based on sparse representation (SPARISH). An overcomplete visual dictionary was first trained based on natural image patches, and image blocks were represented based on the dictionary, producing the sparse coefficients. Then block-wise energy of the image was computed based on the sparse coefficients, which was mainly to account for the attention of high-frequency information caused by blurring. Finally, a variance-based normalization procedure was further developed to achieve consistent sharpness evaluation across different images.

The aforementioned image sharpness metrics have achieved notable success in revealing the influence of blurring on image quality. However, these methods are mainly designed and tested on synthetic blur, which is mainly to simulate real-world camera defocus distortion. This kind of blur is generated by applying the Gaussian low-pass filtering on high quality pristine images, so it is simple and distributed uniformly in the whole image. However, in real-world imaging environments, the real blur is much more complex. The existing image sharpness metrics are very limited in evaluating real blur, which will be demonstrated later in this paper. To design more advanced image sharpness models that can be applied to both synthetic and real blur is still an open problem.

Motivated by the above facts, this paper presents a more practical no-reference and Robust Image Sharpness Evaluation (RISE) metric. The proposed RISE model is inspired by the following two facts. 1) The human visual system (HVS) exhibits multi-scale property when perceiving visual scenes, and hierarchical structures of an image are contained in its scale-space representation [30]–[32]. 2) Viewing distance and accordingly image resolution have great impact on the perceived image quality [33]. For a given image, the scale space representation is first constructed. Then sharpness-aware features are extracted in both the spatial and spectral domains simultaneously. In order to account for the impact of viewing distances on image quality, the image is also down-sampled by several times, producing images at different resolutions. Then the multi-resolution DCT domain entropies are also computed as sharpness features. Finally, all extracted features are utilized to train a Support Vector Regression (SVR) model for sharpness prediction. Extensive experiments are conducted on both synthetically and real blurred image databases. The experimental results confirm the superiority of the proposed metric over the existing sharpness metrics and general-purpose NR image quality metrics. Furthermore, RISE is robust, which means that it has very good generalization ability.

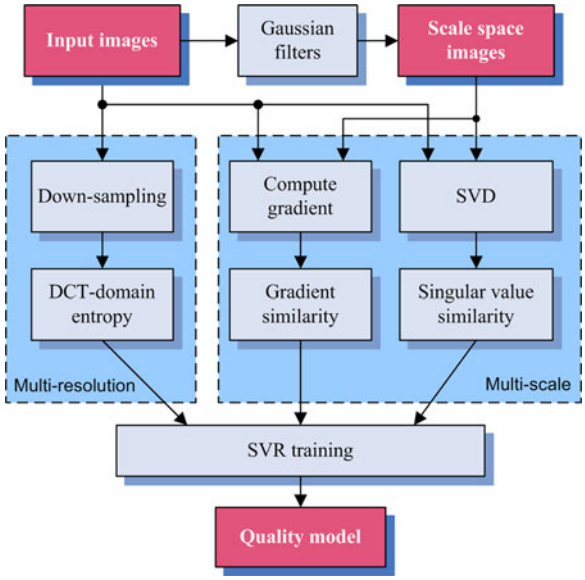

Fig. 1. Flowchart of model training of the proposed RISE metric.

II. Proposed NR Image Sharpness Metric

The HVS is typically a multi-scale device, where the receptive fields have been evolved to have the capacity of extracting hierarchial structures from the scale-space representation of an image [30]. Furthermore, it has been widely accepted that viewing distance has great impact on image quality perception [33]. In accordance with these, the basic idea of the proposed metric is to learn quality-aware features in both multi-scale and multi-resolution domains. Fig. 1 shows the flowchart of the training phase of the proposed metric. For an input image, the scale space representation is first built. Then we extract sharpness-aware features based on the scale space images in gradient domain and Singular Value Decomposition (SVD) domain, respectively. To further account for the impact of viewing distance, we also extract entropy features in several down-sampled versions of the original image. Finally, SVR is employed to train the quality model.

A. Image Scale Space

It has been demonstrated that under various reasonable assumptions the Gaussian function is the only suitable kernel for building image scale space [34]. For an image $I(x, y)$, the scale space $L(x, y, \sigma)$ can be obtained by convoluting it with a set of Gaussian kernels with variable scales

$$L(x, y, \sigma) = I(x, y) * G(x, y, \sigma) \tag{1}$$

where σ is the scale, $*$ denotes the convolution, and the two-dimensional Gaussian function $G(x, y, \sigma)$ is defined as

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2}. \tag{2}$$

In this work, we construct image scale space with five scales, which are denoted by scales 0-4, where scale 0 denotes the original image. Specifically, the Matlab function *imfilter* with Gaussian low-pass filter is adopted to generate the scale space

Fig. 2. Scale space representations of two original images. Original image "Orchid" in the first row is of high quality, while the original image "Woman" in the second row is blurred. (a) Original image. (b) Filter size 3×3, σ=2. (c) Filter size 9×9, σ=4. (d) Filter size 15×15, σ=6. (e) Filter size 21×21, σ=8.

images. The filter sizes are 3×3, 9×9, 15×15 and 21×21, and the corresponding standard deviations (σ) are 2, 4, 6 and 8, respectively.

Fig. 2 shows the scale spaces of two images with different extents of blurring using the aforementioned parameters. It is obvious that the scale space images are in fact reblurred versions of the original ones, and the extents of blur are determined by the Gaussian kernels. It is also observed that for different original images, their scale space images have different characteristics. Specifically, for a high-quality sharp image, the scale space images are quite different, which can be clearly seen from images in the first row. For a blurred image, the corresponding scale space images are very similar, which can be seen from images in the second row. This indicates that the similarities between the scale space images and the original image can be employed to measure the sharpness of the original image. This inspires us to extract sharpness-aware features in image scale space.

B. Multiscale Gradient Similarity

Gradient has been proved to be an effective structure feature for image quality assessment [35]–[37]. Here, we extract gradient similarity feature between each scale space image and the original image. For an original image and the corresponding scale space images, their gradient maps are first computed by

$$\mathbf{D}_i = \frac{|\mathbf{D}_{i,1}| + |\mathbf{D}_{i,2}|}{2} \tag{3}$$

with

$$\mathbf{D}_{i,1} = [-1 \;\; 0 \;\; 1] * \mathbf{L}_i, \quad \mathbf{D}_{i,2} = [-1 \;\; 0 \;\; 1]^{\mathrm{T}} * \mathbf{L}_i \tag{4}$$

where $i = 0, 1, \cdots, 4$, $\mathbf{D}_{i,1}$ and $\mathbf{D}_{i,2}$ denote the horizontal and vertical gradient images in the ith scale, $[-1 \;\; 0 \;\; 1]$ and $[-1 \;\; 0 \;\; 1]^{\mathrm{T}}$ are the horizontal and vertical gradient operators, and $\mathbf{L}_i$ denotes the ith scale space image.

With the gradient maps, we compute the gradient similarity between each scale space image and the original one

$$\mathbf{GS}_k = \frac{2\mathbf{D}_k\mathbf{D}_0 + c_1}{\mathbf{D}_k^2 + \mathbf{D}_0^2 + c_1} \tag{5}$$

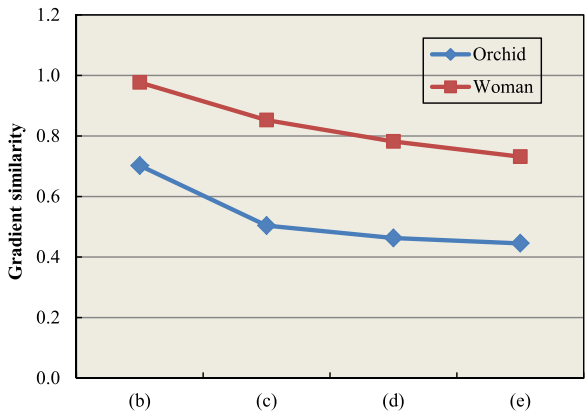

Fig. 3. Multiscale gradient similarity features of images shown in Fig. 2.

where $k = 1, 2, 3, 4$, c_1 is a small constant to ensure numerical stability. Finally, average pooling is employed to generate the final gradient similarity features

$$f_k^G = \frac{1}{MN}\sum_{x=1}^{M}\sum_{y=1}^{N}\mathbf{GS}_k(x, y) \tag{6}$$

where $k = 1, 2, 3, 4$, and $M \times N$ denotes the size of the image.

Fig. 3 shows the curves of the gradient similarity features between the scale space images (b), (c), (d), (e) and the original image (a) in Fig. 2. It is observed that the gradient similarity values of the high-quality image "Orchid" are lower than those of the blurred image "Woman". As a result, the gradient similarity features indicate the extents of blurring in the original images.

C. Multiscale Singular Value Similarity

Besides the gradient similarity features, we also extract multiscale sharpness features in the SVD domain, which can effectively gauge intrinsic structural changes in images [38].

For an $M \times N$ image $\mathbf{A}$, it can be decomposed as

$$\mathbf{A} = \mathbf{USV}^{\mathrm{T}} \tag{7}$$

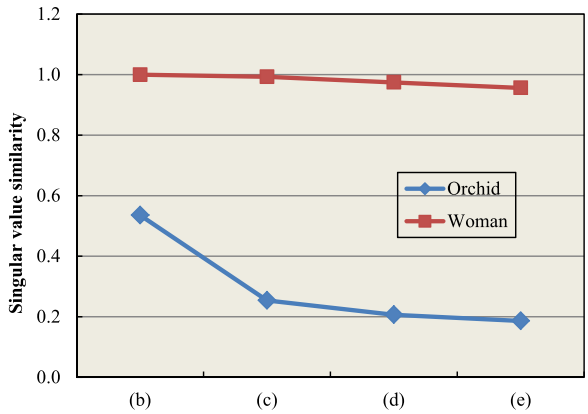

Fig. 4. Multiscale singular value features of images shown in Fig. 2.

where $\mathbf{U}$ is an $M \times M$ unitary matrix satisfying $\mathbf{U}^{\mathrm{T}}\mathbf{U} = \mathbf{I}$, $\mathbf{V}$ is an $N \times N$ unitary matrix satisfying $\mathbf{V}^{\mathrm{T}}\mathbf{V} = \mathbf{I}$, $\mathbf{S}$ is an $M \times N$ matrix with non-negative real numbers (singular values) on the diagonal. The columns of $\mathbf{U}$ and $\mathbf{V}$ are the left and right singular vectors of $\mathbf{A}$, respectively. Let r denote the rank of image matrix $\mathbf{A}$, then the singular value vector can be represented as $\mathbf{s} = (\sigma_1, \sigma_2, \cdots, \sigma_r)$, where $\sigma_i, i = 1, 2, \cdots, r$, is the singular value.

In this work, the singular value vectors of the scale space images are denoted by $\{\mathbf{s}_i, i = 0, 1, 2, 3, 4\}$. Then the singular value similarity between each scale space image and the original image is computed as

$$f_k^S = \frac{2\mathbf{s}_k\mathbf{s}_0 + c_2}{\mathbf{s}_k^2 + \mathbf{s}_0^2 + c_2} \tag{8}$$

where $k = 1, 2, 3, 4$, c_2 is a small constant to avoid numerical instability.

Fig. 4 shows the curves of the singular value similarity features between the scale space images (b), (c), (d), (e) and the original image (a) in Fig. 2. It is observed that the similarity values of the high-quality image "Orchid" are much lower than those of the blurred image "Woman". Furthermore, the differences are more obvious than the gradient similarity curves shown in Fig. 3. Therefore, the singular value similarity features are also effective in portraying blur in images.

D. Multiresolution Entropy in DCT Domain

Entropy measures the amount of information in a signal. For an image, the amount of information reduces when it is subject to blurring. With this consideration, we further extract entropy as quality feature.

Recently, it has been pointed out that the perceived quality of an image is greatly affected by the viewing distance and image resolution [33]. Fig. 5 shows such an example. In this example, the original blurred image (a) is down-sampled by $2\times, 4\times$ and $5\times$ times in both horizontal and vertical directions, producing the images shown in (b)-(d), respectively. It is observed that there is obvious blur in the original image. However, with the reduction of spatial resolution (equivalent of increased viewing distance), the down-sampled images look sharper than the original image.

Inspired by this fact, we first construct a multi-resolution representation of the original image. Then the entropy features are computed in the original and down-sampled images. In this work, we down-sample the original image by $2\times$ and $4\times$ times in both horizontal and vertical directions, producing two lower-resolution images. In implementation, we utilize Matlab function *imresize* with *bicubic* interpolation algorithm to conduct the down-sampling. Bicubic interpolation is an effective interpolation method, which can produce images with high fidelity. By this means, the down-sampling operation does not cause extra distortions. After down-sampling, we then adopt the method in [49] to compute the DCT domain entropy. Specifically, an image is first portioned into non-overlapping 8×8 blocks, which are then transformed into DCT domain, producing the coefficients $C(i, j), i, j = 1, 2, \cdots, 8$. Then the DCT coefficients are normalized as

$$P(i, j) = \frac{C^2(i, j)}{\sum_{i=1}^{8}\sum_{j=1}^{8} C^2(i, j)} \tag{9}$$

where i and j are not equal to 1 simultaneously, so that the DC component is excluded in the normalization. Then the entropy of a local block is calculated by

$$E_k = -\sum_{i=1}^{8}\sum_{j=1}^{8} P(i, j)\log_2 P(i, j) \tag{10}$$

where $k = 1, 2, \cdots, K$, K denotes the total number of blocks in an image. For an image, the entropy values of all blocks are then sorted in descending order, and the mean of the top 40% entropy values is calculated as the final entropy feature. For the two down-sampled images, the same features are computed, so a total of three entropy features are used in this work.

E. Model Training and Sharpness Prediction

In this work, a total of 11 sharpness-aware features are extracted for each image, including 4 multi-scale gradient similarity features, 4 multi-scale singular value similarity features and 3 multi-resolution entropy features. In order to combine them to generate an overall sharpness score, we employ the support vector regression [29] to learn the quality model. Then the trained SVR model is utilized to predict the sharpness score of a query image. In implementation, the Radial Basis Function (RBF) is used as the SVR kernel. Further details of SVR can be found in [29].

III. Experimental Results and Analysis

A. Experiment Settings

In this section, we conduct extensive experiments on public blur image databases to verify the performance of the proposed RISE metric. Specifically, six databases are used in our experiments, including four synthetically and two real blurred image databases. The synthetic blur databases include LIVE [14], CSIQ [39], TID2008 [40], TID2013 [41], while the real

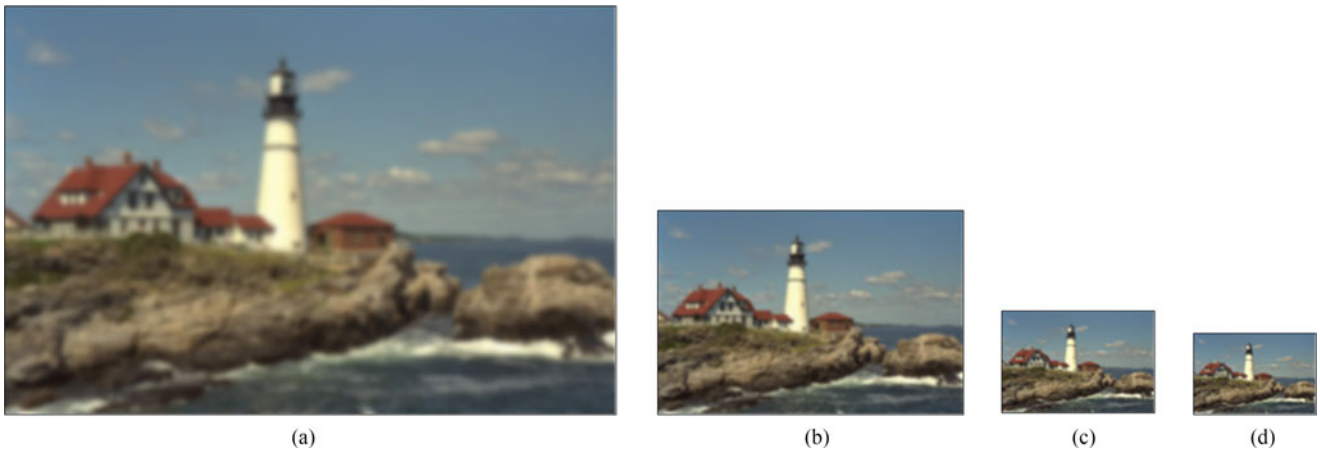

Fig. 5. Image with different resolutions. (a) Original image. (b) Down-sampled image by 2× in both directions. (c) Down-sampled image by 4× in both directions. (d) Down-sampled image by 5× in both directions.

TABLE I
INFORMATION OF BLURRED IMAGE DATABASES USED FOR PERFORMANCE EVALUATION. MOS DENOTES THE MEAN OPINION SCORE, WHILE DMOS DENOTES THE DIFFERENCE MEAN OPINION SCORE

Database	Blur type	No. of images	Subjective score
LIVE	Synthetic	145	DMOS
CSIQ	Synthetic	150	DMOS
TID2008	Synthetic	100	MOS
TID2013	Synthetic	125	MOS
CID2013	Real	473	MOS
BID	Real	586	MOS

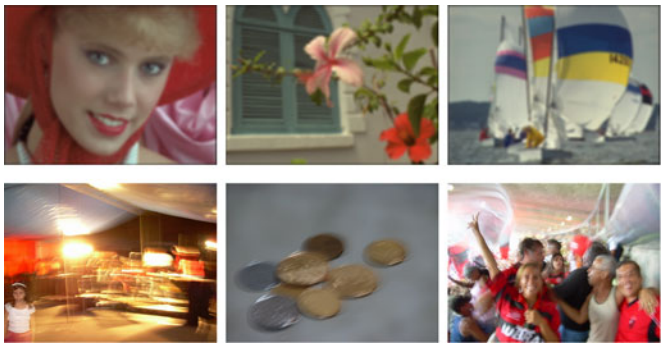

Fig. 6. Sample images. The first row shows the synthetically blurred images, while the second row shows the real blurred images.

blur databases are Camera Image Database (CID2013) [42] and Blurred Image Database (BID) [43]. Detailed information of these databases is summarized in Table I. To be more specific, the synthetic blur images are generated applying Gaussian low-pass filtering on high-quality pristine images, mainly aiming to simulate practical defocus blurring. By contrast, the images in CID2013 and BID are captured by consumer-type cameras in uncontrolled environments. Fig. 6 shows several sample images of both synthetic and real blurring. It is clear that synthetic blur distributes uniformly in the whole image, while real blur is much more complex.

To quantitatively measure the performance of the proposed metric, we adopt three commonly used criteria, including the Pearson linear correlation coefficient (PLCC), root mean square error (RMSE) and the Spearman rank order correlation (SRCC) [14]. PLCC and RMSE are used for measuring prediction accuracy, while SRCC is used for measuring monotonicity. Before calculating these values, a nonlinear fitting is needed to map the objective scores to the same scales of the ground truth subjective scores [44]. In this work, the following five-parameter logistic function is used [14]:

$$f(x) = \tau_1 \left(\frac{1}{2} - \frac{1}{1 + e^{\tau_2 (x - \tau_3)}} \right) + \tau_4 x + \tau_5 \qquad (11)$$

where $\tau_i, i = 1, 2, \cdots, 5$, are the fitting parameters.

B. Performance Evaluation

1) Comparison With NR Sharpness Metrics: In this part, we compare the performance of the proposed RISE metric with ten existing (from classical to the most up to date) no-reference image sharpness metrics, which are Marz. [19], JNB [20], CPBD [21], S3 [22], FISH [23], SVC [24], LPC-SI [25], MLV [26], ARISM [27] and SPARISH [28]. Table II summarizes the experimental results in the six databases, where the best performance values are marked boldfaced. In this experiment, 80% of the images in each database are randomly selected for model training and the remaining 20% images are used for test. To avoid bias, the above training-test operation is conducted 1000 times, and the median performance values are reported for comparison [47]. It should be noted that the performances of the compared sharpness metrics are tested using source codes released by the original authors.

It is known from Table II that the proposed RISE model achieves the best performances in five of the six databases, regardless of prediction accuracy and monotonicity. For synthetic blur, RISE performs the best in CSIQ, TID2008 and TID2013. Particularly in TID2008 and TID2013, RISE significantly outperforms all the other metrics. In LIVE database, the proposed method achieves the best prediction accuracy, while the

TABLE II
PERFORMANCES OF RISE AND THE EXISTING SHARPNESS/BLUR MODELS ON FOUR SYNTHETIC AND TWO REAL BLUR DATABASES

Dataset	Criterion	Marz. [19]	JNB [20]	CPBD [21]	S3 [22]	FISH [23]	SVC [24]	LPC-SI [25]	MLV [26]	ARISM [27]	SPARISH [28]	RISE
LIVE	PLCC	0.8102	0.8188	0.9124	0.9526	0.9044	0.9384	0.9316	0.9429	0.9560	0.9596	**0.9620**
	SRCC	0.8008	0.7874	0.9189	0.9436	0.8856	0.9257	0.9389	0.9312	0.9511	**0.9598**	0.9493
	RMSE	10.8261	10.6027	7.5583	5.6204	7.8800	6.3812	6.7116	6.1521	5.4167	5.1958	**5.0011**
CSIQ	PLCC	0.7273	0.8711	0.8282	0.9175	0.8729	0.9322	0.9322	0.9444	0.9410	0.9391	**0.9463**
	SRCC	0.7661	0.8380	0.8847	0.9059	0.8941	0.9019	0.9071	0.9247	0.9261	0.9141	**0.9279**
	RMSE	0.1967	0.1407	0.1606	0.1140	0.1398	0.1037	0.1037	0.0943	0.0970	0.0984	**0.0926**
TID2008	PLCC	0.7111	0.6961	0.8331	0.8551	0.8138	0.8469	0.8635	0.8593	0.8430	0.8955	**0.9289**
	SRCC	0.7283	0.6685	0.8412	0.8643	0.7972	0.8384	0.8736	0.8548	0.8505	0.8963	**0.9218**
	RMSE	0.8250	0.8424	0.6492	0.6084	0.6820	0.6241	0.5918	0.6001	0.6312	0.5224	**0.4278**
TID2013	PLCC	0.7708	0.7115	0.8620	0.8816	0.8333	0.8752	0.8940	0.8830	0.8954	0.9020	**0.9419**
	SRCC	0.7620	0.6947	0.8531	0.8609	0.8063	0.8612	0.8984	0.8787	0.8982	0.8940	**0.9338**
	RMSE	0.7951	0.8769	0.6325	0.5890	0.6899	0.6036	0.5591	0.5858	0.5556	0.5389	**0.4201**
CID2013	PLCC	0.5247	0.5373	0.5245	0.6863	0.6521	0.3362	0.7031	0.6890	0.5523	0.6775	**0.7934**
	SRCC	0.4394	0.4511	0.4448	0.6460	0.5893	0.2501	0.6024	0.6206	0.4719	0.6607	**0.7690**
	RMSE	19.4510	19.2699	19.4530	16.6190	17.3225	21.5182	16.2474	16.5594	19.0475	16.8056	**13.7823**
BID	PLCC	0.2636	0.2612	0.2704	0.4271	0.4799	0.4287	0.3901	0.3643	0.1841	0.3460	**0.6017**
	SRCC	0.2515	0.2383	0.2717	0.4253	0.4882	0.3573	0.3161	0.3236	0.1742	0.3413	**0.5839**
	RMSE	1.2077	1.2085	1.2053	1.1320	1.0983	1.1310	1.1528	1.1659	1.2305	1.1746	**0.9936**

TABLE III
SUMMARY OF STATISTICAL PERFORMANCES BETWEEN RISE AND THE EXISTING NR IMAGE SHARPNESS METRICS IN SIX DATABASES. THE VALUE 1 (0) INDICATES THAT RISE PERFORMS STATISTICALLY BETTER (COMPETITIVE) THAN THE COMPARED METRIC

Metric	LIVE	CSIQ	TID2008	TID2013	CID2013	BID
Marz. [19]	1	1	1	1	1	1
JNB [20]	1	1	1	1	1	1
CPBD [21]	1	1	1	1	1	1
S3 [22]	0	1	1	1	1	1
FISH [23]	1	1	1	1	1	1
SVC [24]	1	0	1	1	1	1
LPC [25]	1	0	1	1	1	1
MLV [26]	1	0	1	1	1	1
ARISM [27]	0	0	1	1	1	1
SPARISH [28]	0	0	1	1	1	1

prediction monotonicity ranks the third. For real blur, the proposed method also delivers significantly better performances in both CID2013 and BID databases. In CID2013, RISE achieves 0.7934 and 0.7690 for PLCC and SRCC, while the second best results are only PLCC= 0.7031 (delivered by LPC-SI [25]) and SRCC = 0.6607 (delivered by SPARISH [28]), respectively. In BID, our metric achieves 0.6017 and 0.5839, while the second best results (delivered by FISH [23]) are only 0.4799 and 0.4882. From these results, we know that the proposed metric outperforms the existing sharpness metrics for both synthetic and real blur.

In order to have an intuitive understanding of the statistical significance of the proposed RISE metric in relative to the existing NR image sharpness metrics, we further conduct the F-test following the same method in [28]. In this experiment, 95% confidence level is adopted for the F-test. Table III summarizes the simulation results, where value 1 (0) indicates that RISE performs statistically better (competitive) than the compared metric. It is observed from the table that the proposed metric performs statistically better than all the existing sharpness metrics in TID2008, TID2013, CID2013 and BID databases. In LIVE, only S3 [22], ARISM [27] and SPARISH [28] are comparable to the proposed metric. In CSIQ, RISE performs statistically better than five metrics, and the remaining five metrics are statistically comparable to the proposed RISE. From these results, we know that the overall statistical performance of the proposed metric is the best.

2) Comparison With General-Purpose NR Image Quality Metrics: In the literature, several general-purpose NR image quality metrics have been proposed, which can predict image quality without knowing specific distortion types. As a metric specifically designed for image sharpness, the proposed method is expected to outperform these metrics. To this end, we further compare RISE with the state-of-the-art general-purpose NR image quality models, including BIQI [45], BLLINDS-II [46], BRISQUE [47], DIIVINE [48], SSEQ [49], NIQE [50] and QAC [51]. The source codes of these metrics are provided by the original authors. In this experiment, most of the compared metrics are learning-based methods, and their models are trained using the LIVE database. For fairness, we also train our model using the LIVE database. In other words, for synthetic blurring, all models are first trained in LIVE database, and then they are used for sharpness evaluation in the other three databases, i.e., CSIQ, TID2008 and TID2013. For CID2013 and BID, since models trained using synthetically blurred images in LIVE are not applicable to real blurred images, we re-train all models (both RISE and the compared metrics) following the method described in the above subsection. Namely, 80% of the images are used for model training, and the remaining 20% are used for test. This process is also repeated 1000 times and the median values are reported. Table IV lists the experimental results, where the best performances are marked boldfaced.

TABLE IV
PERFORMANCES OF RISE AND THE GENERAL-PURPOSE NO-REFERENCE QUALITY METRICS ON THREE SYNTHETIC AND TWO REAL BLUR DATABASES

Database	Criterion	BIQI [45]	BLIINDS-II [46]	BRISQUE [47]	DIIVINE [48]	SSEQ [49]	NIQE [50]	QAC [51]	RISE
CSIQ	PLCC	0.8470	0.8870	0.9249	0.8824	0.8711	0.9265	0.8352	**0.9555**
	SRCC	0.7713	0.8915	0.9025	0.8716	0.8732	0.8944	0.8362	**0.9389**
	RMSE	0.1523	0.1323	0.1090	0.1348	0.1407	0.1078	0.1576	**0.0845**
TID2008	PLCC	0.7940	0.8419	0.8656	0.8350	0.8578	0.8323	0.8134	**0.9167**
	SRCC	0.7988	0.8589	0.8648	0.8294	0.8521	0.8233	0.8116	**0.9138**
	RMSE	0.7134	0.6332	0.5876	0.6457	0.6031	0.6505	0.6827	**0.4689**
TID2013	PLCC	0.8209	0.8583	0.8699	0.8479	0.8668	0.8148	0.8481	**0.9285**
	SRCC	0.8186	0.8626	0.8666	0.8440	0.8633	0.8035	0.8466	**0.9339**
	RMSE	0.7127	0.6403	0.6155	0.6616	0.6222	0.7234	0.6613	**0.4635**
CID2013	PLCC	0.7702	0.7250	0.7040	0.5219	0.6761	0.6885	0.1776	**0.7934**
	SRCC	0.7426	0.6972	0.6784	0.4451	0.6605	0.6292	0.1584	**0.7690**
	RMSE	14.4158	15.4924	15.9511	19.2129	16.6101	16.5699	22.4854	**13.7823**
BID	PLCC	0.6003	0.5466	**0.6101**	0.4786	0.5996	0.4618	0.3197	0.6017
	SRCC	0.5750	0.5284	**0.5857**	0.4341	0.5780	0.4598	0.3084	0.5839
	RMSE	0.9951	1.0443	**0.9859**	1.0903	0.9924	1.1104	1.1862	0.9936

TABLE V
SUMMARY OF STATISTICAL PERFORMANCES BETWEEN RISE AND THE EXISTING GENERAL-PURPOSE NR IMAGE QUALITY METRICS IN FIVE DATABASES. THE VALUE 1 (0) INDICATES THAT RISE PERFORMS STATISTICALLY BETTER (COMPETITIVE) THAN THE COMPARED METRIC

Metric	CSIQ	TID2008	TID2013	CID2013	BID
BIQI [45]	1	1	1	0	0
BLIINDS-II [46]	1	1	1	1	0
BRISQUE [47]	1	1	1	0	0
DIIVINE [48]	1	1	1	1	1
SSEQ [49]	1	1	1	1	0
NIQE [50]	1	1	1	1	1
QAC [51]	1	1	1	1	1

It is observed from Table IV that the proposed RISE model significantly outperforms the general-purpose NR image quality metrics in all synthetic blur databases. In real blur database CID2013, the proposed metric produces the best results. In BID, our metric ranks the second, which is only slightly worse than the best performing BRISQUE [47].

Similar to Table III, F-test is also conducted to know the statistical significance of the proposed metric in relative to the existing general-purpose NR image quality metrics. The experimental results are summarized in Table V. It is easily observed from the table that the proposed metric performs statistically better than all the compared general-purpose NR image quality metrics in CSIQ, TID2008 and TID2013 databases. In CID2013, only BIQI [45] and BRISQUE [47] are statistically competitive to the proposed RISE metric. In BID database, the proposed metric performs statistically better than DIIVINE [48], NIQE [50] and QAC [51]. The performances of the remaining metrics are statistically competitive. Therefore, the proposed metric achieves the best overall performance.

From Tables II to V, we have the following observations. 1) The current image sharpness metrics are mainly effective for synthetic blur, but they are quite limited in predicting real blur. 2) Compared to the sharpness metrics, the general-purpose image quality metrics are more effective in evaluating real blur than synthetic blur. 3) The proposed RISE metric is effective for both synthetic and real blur, and it delivers the best overall performance for both types of blur.

Fig. 7 shows several images with either synthetic or real blur, together with their subjective (MOS) and predicted RISE scores. It is observed from the results that with the increase of MOS values from (a) to (h) (high MOS value corresponds to high quality), the predicted sharpness scores also increase accordingly. This indicates that the predicted sharpness scores are consistent with the subjective ratings.

C. Generalization Ability

For learning-based quality metrics, generalization ability is an important issue. In this part, we test the generalization ability of the proposed metric using cross-validation. Specifically, we conduct two types of cross-validations. Type I: For synthetic blur, we train our model in one of the four databases, and then we use the trained model for sharpness evaluation on the other three databases. Type II: For real blur, we train our models in CID2013 and BID, and then we use the trained models for sharpness evaluation on the four synthetic blur databases. Since real blur is much more complex than synthetic blur, training a quality model in synthetic blur and then testing in real blur is not expected to perform well. So we do not include this kind of cross-validation.

Table VI lists the simulation results of type I cross-validation. Four models are trained using the four synthetic blur databases respectively, and the trained models are used for sharpness evaluation of the other three databases. From the results, we know that for the four databases, very encouraging results have been achieved. Most of the performance values are higher than 0.85, both for PLCC and SRCC. The best results are achieved when the model is trained in LIVE database, where the PLCC and SRCC values are all above 0.91.

Table VII lists the simulation results of type II cross-validation. In this experiment, two models are trained using

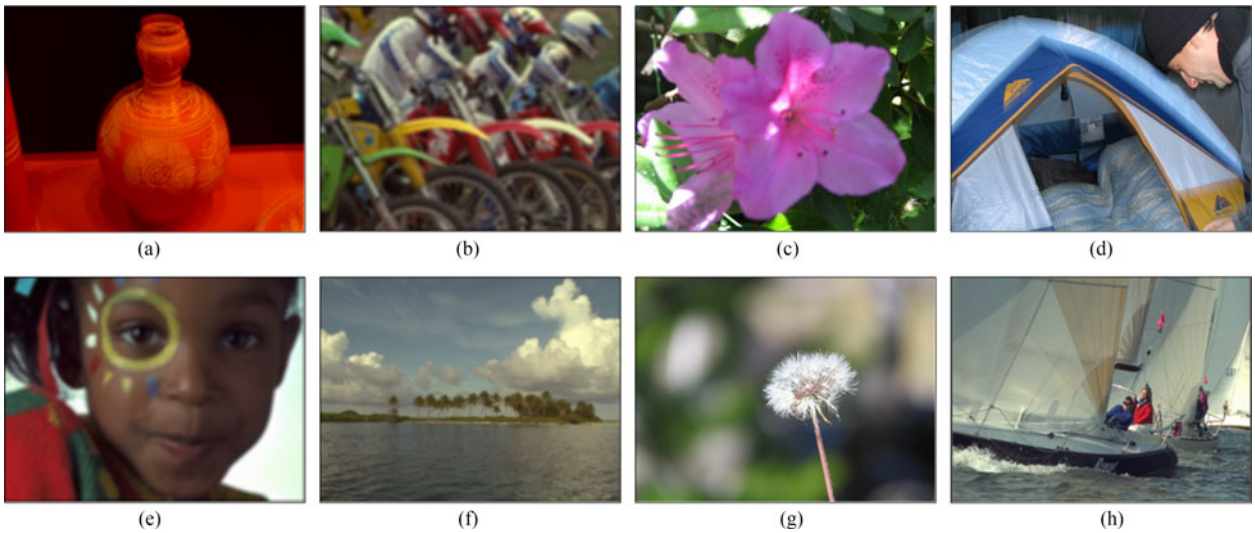

Fig. 7. Sample blurred images, together with their subjective and predicted RISE scores. Images (a), (c), (d), and (g) are real blurred images from the BID database; images (b), (e), (f), and (h) are synthetically blurred images from the TID2008 and TID2013 databases. (a) MOS=0.8280, RISE=1.0150. (b) MOS=1.2500, RISE=1.3506. (c) MOS=1.4239, RISE=2.3972. (d) MOS=2.8391, RISE=3.0765. (e) MOS=3.3636, RISE=3.2060. (f) MOS=3.4000, RISE=3.3440. (g) MOS=3.9971, RISE=3.4299. (h) MOS=4.3514. RISE=4.5315.

TABLE VI
CROSS-VALIDATION OF THE PROPOSED RISE METRIC FOR SYNTHETIC BLUR (TYPE I)

Training database	Test database			
	Criterion	CSIQ	TID2008	TID2013
LIVE	PLCC	0.9555	0.9167	0.9285
	SRCC	0.9389	0.9138	0.9339
	RMSE	0.0845	0.4689	0.4635
	Criterion	LIVE	TID2008	TID2013
CSIQ	PLCC	0.9305	0.8509	0.8834
	SRCC	0.9119	0.8538	0.8930
	RMSE	6.7665	0.6165	0.5848
	Criterion	LIVE	CSIQ	TID2013
TID2008	PLCC	0.8802	0.8945	0.8880
	SRCC	0.8638	0.8495	0.8696
	RMSE	8.7675	0.1281	0.5738
	Criterion	LIVE	CSIQ	TID2008
TID2013	PLCC	0.8149	0.8938	0.9734
	SRCC	0.8497	0.8913	0.9764
	RMSE	10.7071	0.1285	0.2687

real blur databases CID2013 and BID. Then the trained models are used for sharpness evaluation in the four synthetic blur databases. It is clear that the experimental results are very encouraging. Most of the performance values are higher than 0.80, which indicates that the predicted sharpness scores are very consistent with the ground truth subjective scores.

From these results, we know that the proposed metric is very robust because it has very good generalization ability, which is highly desired in real-world applications.

D. Impact of Scales

The proposed metric is based on sharpness feature extraction in image scale space. Here, we investigate the impact of scale number on the metric performance. Specifically, we test the performance of the proposed RISE metric with different number of scales, ranging from 2 to 5. Table VIII summarizes the experimental results in the six databases, where the best results in each database are marked in boldface. In order to know the overall performance, the database-size weighted average results are also provided.

TABLE VII
PERFORMANCES OF THE PROPOSED METHOD WHEN THE MODELS TRAINED IN CID2013 AND BID ARE USED FOR SHARPNESS EVALUATION IN LIVE, CSIQ, TID2008, AND TID2013 (TYPE II)

Training database	Test database				
	Criterion	LIVE	CSIQ	TID2008	TID2013
CID2013	PLCC	0.8660	0.8374	0.8033	0.8235
	SRCC	0.8040	0.7430	0.7929	0.8101
	RMSE	9.2375	0.1566	0.6990	0.7079
BID	PLCC	0.9039	0.8875	0.8031	0.7718
	SRCC	0.8760	0.7906	0.8017	0.7160
	RMSE	7.9018	0.1321	0.6992	0.7937

From Table VIII, we have the following observations. When only 2 scales are used, the performances are not satisfactory, especially for real blurred images. Therefore, multiple scales are needed for effective sharpness evaluation. For the four synthetic blur databases, namely LIVE, CSIQ, TID2008 and TID2013, 3 to 4 scales are sufficient for obtaining very good performances. However for the two real blur databases CID2013 and BID, the best performances are achieved when 5 scales are used. This indicates that real blur is more complex and needs feature representations in more scales. The best overall performance is obtained when 5 scales are used. It should be noted that 3 or 4 scales can also be used in this paper, which reduces the number of features with the sacrifice of slightly compromised performance for real blurred images.

TABLE VIII
IMPACT OF SCALE NUMBER ON THE PERFORMANCES OF THE PROPOSED METRIC

No. of scales	Criterion	LIVE	CSIQ	TID2008	TID2013	CID2013	BID	Weighted average
2	PLCC	0.9519	0.9484	0.9261	0.9203	0.6953	0.5281	0.7133
	SRCC	0.9359	0.9341	0.9157	0.9077	0.6569	0.5157	0.6927
3	PLCC	0.9621	**0.9507**	0.9379	**0.9511**	0.7806	0.5948	0.7679
	SRCC	0.9493	**0.9344**	0.9293	**0.9442**	0.7527	0.5786	0.7497
4	PLCC	**0.9630**	0.9490	**0.9408**	0.9500	0.7886	0.5961	0.7708
	SRCC	**0.9533**	0.9287	**0.9310**	0.9430	0.7622	0.5790	0.7526
5	PLCC	0.9620	0.9463	0.9289	0.9419	**0.7934**	**0.6017**	**0.7726**
	SRCC	0.9493	0.9279	0.9218	0.9338	**0.7690**	**0.5839**	**0.7547**

TABLE IX
PERFORMANCES OF THE PROPOSED METRIC WHEN DIFFERENT PERCENTAGES OF IMAGES ARE USED FOR MODEL TRAINING

Traning-test	Criterion	LIVE	CSIQ	TID2008	TID2013	CID2013	BID
80%-20%	PLCC	0.9620	0.9463	0.9289	0.9419	0.7934	0.6017
	SRCC	0.9493	0.9279	0.9218	0.9338	0.7690	0.5839
60%-40%	PLCC	0.9483	0.9320	0.8981	0.9122	0.7633	0.5773
	SRCC	0.9376	0.9173	0.8947	0.9066	0.7456	0.5652
50%-50%	PLCC	0.9449	0.9281	0.8810	0.8959	0.7486	0.5656
	SRCC	0.9346	0.9117	0.8777	0.8878	0.7317	0.5551
40%-60%	PLCC	0.9412	0.9232	0.8627	0.8769	0.7276	0.5469
	SRCC	0.9300	0.9043	0.8607	0.8698	0.7114	0.5370

TABLE X
CONTRIBUTIONS OF THE THREE COMPONENTS IN THE PROPOSED MODEL

Feature	Criterion	LIVE	CSIQ	TID2008	TID2013	CID2013	BID
Gradient	PLCC	0.9425	0.9399	0.9423	0.9526	0.5418	0.3732
	SRCC	0.9301	0.9217	0.9338	0.9431	0.5290	0.3541
SVD	PLCC	0.9569	0.9628	0.9211	0.9364	0.6933	0.4786
	SRCC	0.9453	0.9379	0.8951	0.9231	0.6253	0.4629
Entropy	PLCC	0.9276	0.9334	0.9115	0.9132	0.5824	0.4070
	SRCC	0.9046	0.9043	0.9053	0.9018	0.5128	0.3827
All	PLCC	0.9620	0.9463	0.9289	0.9419	0.7934	0.6017
	SRCC	0.9493	0.9279	0.9218	0.9338	0.7690	0.5839

E. Impact of Training Images

In order to have an intuitive understanding of how the number of training images affect the metric performance, we also conduct an experiment to test the proposed method when different numbers of images are used for model training. In this experiment, we use four different percentages of images for model training, namely 80%, 60%, 50%, 40%. The experimental results are summarized in Table IX. It is clearly known from the table that with the decreasing number of training images, the performance values drop slightly. For the proposed method, even if only 40% images are used for training, the performance is still very good. This is helpful in real-world applications because relatively small amount of images can be used for model training, which can still achieve very satisfactory results.

F. Contributions of Components

In the proposed metric, three kinds of features are extracted to train the quality model, including gradient, SVD and entropy. In this part, we further investigate their individual contributions to the overall performance. To this end, we conduct model training and test using the three groups of features, separately. As aforementioned, this operation is conducted 1000 times and the median PLCC and SRCC values are reported. Table X summarizes the experimental results.

From Table X, we have the following findings. For synthetic blur, each feature can achieve very promising results, and almost all the PLCC and SRCC values are above 0.90. However, neither of them can achieve the best results in all databases. Specifically, gradient feature delivers the best performances in TID2008 and TID2013, while SVD feature performs the best in CSIQ. For real blur, a single feature does not perform well so the individual performances are much lower than that produced by using all three features simultaneously. This is not hard to understand in that the quality of real blurred images are commonly determined by a various of factors, including camera out-of-focus, camera/object motion, image local contrast, etc. As a result, more diversified features are needed to build a more reliable model for real blur. From these results, we know that all three features are needed to achieve consistently good performances for sharpness evaluation, especially for real blurred images captured by consumer-type cameras in uncontrolled environments.

IV. CONCLUSION

In this paper, we have presented a novel objective no-reference quality metric for robust image sharpness evaluation. The proposed method has been inspired by the fact that the human visual system is highly adapted to extract multi-scale and multi-resolution features for scene understanding. With these considerations, we have extracted multi-scale and multi-resolution quality features in both the spatial and spectral domains to learn a sharpness quality model. We have tested the performances of the proposed metric in six blurred image databases, including both synthetic and real blur. The experimental results have confirmed that our metric is advantageous over the existing sharpness and general-purpose image quality metrics. Furthermore, the proposed method has very good generalization ability, because a model trained in one database can be readily applicable to other databases with very good performances.

Based on this study, we have also found that while the state-of-the-art sharpness metrics are effective for synthetic blur, they are quite limited in predicting real blur. The proposed metric is more effective for real blur than the existing metrics, but its performance is far from ideal. To develop more advanced quality models for real blur evaluation is still highly desired.

REFERENCES

[1] A. Hameed, R. Dai, and B. Balas, "A decision-tree-based perceptual video quality prediction model and its application in FEC for wireless multimedia communications," *IEEE Trans. Multimedia*, vol. 18, no. 4, pp. 764–774, Apr. 2016.

[2] S. Jang and J. S. Lee, "On evaluating perceptual quality of online user-generated videos," *IEEE Trans. Multimedia*, vol. 18, no. 9, pp. 1808–1818, Sep. 2016.

[3] K. Gu *et al.*, "Saliency-guided quality assessment of screen content images," *IEEE Trans. Multimedia*, vol. 18, no. 6, pp. 1098–1110, Jun. 2016.

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

[5] D. M. Chandler, "Seven challenges in image quality assessment: Past, present, and future research," *ISRN Signal Process.*, vol. 2013, pp. 1–53, 2013, Art. no. 905685.

[6] S. Q. Wang *et al.*, "Utility-driven adaptive preprocessing for screen content video compression," *IEEE Trans. Multimedia*, 2016, to be published. doi: 10.1109/TMM.2016.2625276.

[7] Z. Q. Pan, J. J. Lei, Y. Zhang, X. M. Sun, and S. Kwong, "Fast motion estimation based on content property for low-complexity H.265HEVC encoder," *IEEE Trans. Broadcast.*, vol. 62, no. 3, pp. 675–684, Sep. 2016.

[8] Z. Q. Pan, Y. Zhang, and S. Kwong, "Efficient motion and disparity estimation optimization for low complexity multiview video coding," *IEEE Trans. Broadcast.*, vol. 61, no. 2, pp. 166–176, Jun. 2015.

[9] S. Q. Wang *et al.*, "Guided image contrast enhancement based on retrieved images in cloud," *IEEE Trans. Multimedia*, vol. 18, no. 2, pp. 219–232, Feb. 2016.

[10] L. D. Li *et al.*, "No-reference quality assessment of enhanced images," *China Commun.*, vol. 13, no. 9, pp. 121–130, Sep. 2016.

[11] L. D. Li *et al.*, "Perceptual quality evaluation for image defocus deblurring," *Sig. Proc., Image Comm.*, vol. 48, pp. 81–91, Oct. 2016.

[12] Z. L. Zhou, Y. L. Wang, Q. M. J. Wu, C. N Yang, and X. M. Sun, "Effective and efficient global context verification for image copy detection," *IEEE Trans. Inf. Forensics Security*, vol. 12, no. 1, pp. 48–63, Jan. 2017.

[13] J. Li, X. L. Li, B. Yang, and X. M. Sun, "Segmentation-based image copy-move forgery detection scheme," *IEEE Trans. Inf. Forensics Security*, vol. 10, no. 3, pp. 507–518, Mar. 2015.

[14] H. R. Sheikh, M. F. Sabir, and A. C. Bovik, "A statistical evaluation of recent full reference image quality assessment algorithms," *IEEE Trans. Image Process.*, vol. 15, no. 11, pp. 3440–3451, Nov. 2006.

[15] F. Qi, D. B. Zhao, and W. Gao, "Reduced reference stereoscopic image quality assessment based on binocular perceptual information," *IEEE Trans. Multimedia*, vol. 17, no. 12, pp. 2338–2344, Dec. 2015.

[16] J. J. Wu, W. S. Lin, G. M. Shi, and A. M. Liu, "Reduced-reference image quality assessment with visual information fidelity," *IEEE Trans. Multimedia*, vol. 12, no. 7, pp. 1700–1705, Nov. 2013.

[17] W. J. Zhou and L. Yu, "Binocular responses for no-reference 3D image quality assessment," *IEEE Trans. Multimedia*, vol. 18, no. 6, pp. 1077–1084, Jun. 2016.

[18] K. Gu. S. Q. Wang *et al.*, "Blind quality assessment of tone-mapped images via analysis of information, naturalness, and structure," *IEEE Trans. Multimedia*, vol. 18, no. 3, pp. 432–443, Mar. 2016.

[19] P. Marziliano, F. Dufaux, S. Winkler, and T. Ebrahimi, "Perceptual blur and ringing metrics: Application to JPEG2000," *Signal Process., Image Commun.*, vol. 19, no. 2, pp. 163–172, Feb. 2004.

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

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

[22] C. T. Vu, T. D. Phan, and D. M. Chandler, "S3: A spectral and spatial measure of local perceived sharpness in natural images," *IEEE Trans. Image Process.*, vol. 21, no. 3, pp. 934–945, Mar. 2012.

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

[24] Q. B. Sang, H. X. Qi, X. J. Wu, C. F. Li, and A. C. Bovik, "No-reference image blur index based on singular value curve," *J. Vis. Commun. Image Represent.*, vol. 25, no. 7, pp. 1625–1630, Oct. 2014.

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

[26] K. Bahrami and A. C. Kot, "A fast approach for no-reference image sharpness assessment based on maximum local variation," *IEEE Signal Process. Lett.* vol. 21, no. 6, pp. 751–755, Jun. 2014.

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

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

[29] C. C. Chang and C. J. Lin, "LIBSVM: A library for support vector machines," *ACM Trans. Intell. Syst. Technol.*, vol. 2, no. 3, pp. 1–27, 2011.

[30] B. M. ter Haar Romeny, *Front-End Vision and Multi-Scale Image Analysis-Multi-Scale Computer Vision Theory and Applications, Written in Mathematics* (Comput. Imaging Vis. Ser.). Berlin, Germany: Springer-Verlag, 2003, pp. 1–466.

[31] K. Gu *et al.*, "Analysis of distortion distribution for pooling in image quality prediction," *IEEE Trans. Broadcast.*, vol. 62, no. 2, pp. 446–456, Jun. 2016.

[32] S. C. Nercessian, K. A. Panetta, and S. S. Agaian, "Non-linear direct multi-scale image enhancement based on the luminance and contrast masking characteristics of the human visual system," *IEEE Trans. Image Process.*, vol. 22, no. 9, pp. 3549–3561, Sep. 2013.

[33] K. Gu, M. Liu, G. T. Zhai, X. K. Yang, and W. J. Zhang, "Quality assessment considering viewing distance and image resolution," *IEEE Trans. Broadcast.*, vol. 61, no. 3, pp. 520–531, Sep. 2015.

[34] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vis.*, vol. 60, no. 2, pp. 91–110, 2004.

[35] A. M. Liu, W. S. Lin, and M. Narwaria, "Image quality assessment based on gradient similarity," *IEEE Trans. Image Process.*, vol. 21, no. 4, pp. 1500–1512, Apr. 2012.

[36] W. F. Xue, L. Zhang, X. Q. Mou, and A. C. Bovik, "Gradient magnitude similarity deviation: A highly efficient perceptual image quality index," *IEEE Trans. Image Process.*, vol. 23, no. 2, pp. 684–695, Feb. 2014.

[37] L. D. Li *et al.*, "Sparse representation-based image quality index with adaptive sub-dictionaries," *IEEE Trans. Image Process.*, vol. 25, no. 8, pp. 3775–3786, Aug. 2016.

[38] M. Narwaria and W. S. Lin, "SVD-based quality metric for image and video using machine learning," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 42, no. 2, pp. 347–364, Apr. 2012.

[39] E. C. Larson and D. M. Chandler, "Most apparent distortion: Full-reference image quality assessment and the role of strategy," *J. Electron. Imaging*, vol. 19, no. 1, pp. 00100:61–001006:21, Mar. 2010 .

[40] N. Ponomarenko *et al.*, "TID2008-a database for evaluation of full-reference visual quality assessment metrics," *Adv. Mod. Radioelectron.*, vol. 10, no. 4, pp. 30–45, 2009.

[41] N. Ponomarenko *et al.*, "Color image database TID2013: Peculiarities and preliminary results," in *Proc. Eur. Workshop Vis. Inf. Process.*, 2013, pp. 106–111.

[42] T. Virtanen, M. Nuutinen, M. Vaahteranoksa, P. Oittinen, and J. Häkkinen, "CID2013: A database for evaluating no-reference image quality assessment algorithms," *IEEE Trans. Image Process.*, vol. 24, no. 1, pp. 390–402, Jan. 2015.

[43] A. Ciancio *et al.*, "No-reference blur assessment of digital pictures based on multifeature classifiers," *IEEE Trans. Image Process.*, vol. 20, no. 1, pp. 64–75, Jan. 2011.

[44] "Final report from the Video Quality Experts Group on the validation of objective models of video quality assessment, phase II," Aug. 2003, accessed on Aug. 2016. [Online]. Available: http://www.vqeg.org

[45] A. K. Moorthy and A. C. Bovik, "A two-step framework for constructing blind image quality indices," *IEEE Signal Process. Lett.*, vol. 17, no. 5, pp. 513–516, May 2010.
[46] M. A Saad and A. C. Bovik, "Blind image quality assessment: A natural scene statistics approach in the DCT domain," *IEEE Trans. Image Process.*, vol. 21, no. 8, pp. 3339–3352, Aug. 2012.
[47] A. Mittal, A. K. Moorthy, and A. C. Bovik, "No-reference image quality assessment in the spatial domain," *IEEE Trans. Image Process.*, vol. 21, no. 12, pp. 4695–4708, Dec. 2012.
[48] A. K. oorthy and A. C. Bovik, "Blind image quality assessment: From natural scene statistics to perceptual quality," *IEEE Trans. Image Process.*, vol. 20, no. 12, pp. 3350–3364, Dec. 2011.
[49] L. X. Liu, B. Liu, H. Huang, and A. C. Bovik, "No-reference image quality assessment based on spatial and spectral entropies," *Signal Process., Image Commun.*, vol. 29, no. 8, pp. 856–863, 2014.
[50] A. Mittal, R. Soundararajan, and A. C. Bovik, "Making a completely blind image quality analyzer," *IEEE Signal Process. Lett.*, vol. 20, no. 3, pp. 209–212, Mar. 2013.
[51] W. F. Xue, L. Zhang, and X. Q. Mou, "Learning without human scores for blind image quality assessment," in *Proc. Int. Conf. Comput. Vis. Pattern Recog.*, 2013, pp. 995–1002.

Leida Li received the B.S. and Ph.D. degrees from Xidian University, Xi'an, China, in 2004 and 2009, respectively.

From February to June 2008, he was a visiting Ph.D. student with the Department of Electronic Engineering, National Kaohsiung University of Applied Sciences, Taiwan, China. From January 2014 to January 2015, he was a Visiting Research Fellow with the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore. He is currently a Professor with the School of Information and Electrical Engineering, China University of Mining and Technology, Xuzhou, China. His research interests include multimedia quality assessment, information hiding, and image forensics.

Wenhan Xia received the B.E. degree from Wuhan Huaxia University of Technology, Wuhan, China, in 2014, and is currently working toward the M.S. degree at the School of Information and Electrical Engineering, China University of Mining and Technology, Xuzhou, China.

Her research interests include image quality assessment and sparse representation.

Weisi Lin (S'91–M'92–SM'00–F'16) received the Ph.D. degree from Kings College, London University, London, U.K., in 1993.

He is currently an Associate Professor with the School of Computer Engineering, Nanyang Technological University, Singapore, and served as a Laboratory Head of Visual Processing, Institute for Infocomm Research, Singapore. He has authored more than 300 scholarly publications, holds 7 patents, and received over $4 million in research grant funding. He has maintained active long-term working relationship with a number of companies. His research interests include image processing, video compression, perceptual visual and audio modeling, computer vision, and multimedia communication.

Prof. Lin is a Chartered Engineer in the U.K., a Fellow of the Institution of Engineering Technology, and an Honorary Fellow in the Singapore Institute of Engineering Technologists. He served as an Associate Editor of the IEEE TRANSACTIONS ON MULTIMEDIA, the IEEE SIGNAL PROCESSING LETTERS, and the *Journal of Visual Communication and Image Representation*. He is also on six IEEE Technical Committees and Technical Program Committees of a number of international conferences. He was the Lead Guest Editor for a Special Issue on perceptual signal processing of the IEEE JOURNAL OF SELECTED TOPICS IN SIGNAL PROCESSING in 2012. He Co-Chaired the IEEE MMTC special interest group on quality of experience. He was an Elected Distinguished Lecturer of APSIPA in 2012/13.

Yuming Fang received B.E. degree from Sichuan University, Sichuan, China, in 2006, the M.S. degree from Beijing University of Technology, Beijing, China, in 2009, and the Ph.D. degree from Nanyang Technological University, Singapore, in 2013.

He is currently an Associate Professor with the School of Information Technology, Jiangxi University of Finance and Economics, Nanchang, China. He was previously a Visiting Researcher with the IRCCyN Laboratory, PolyTech' Nantes, University Nantes, France; National Tsinghua University, Taiwan, China, and the University of Waterloo, Waterloo, ON, Canada. His research interests include visual attention modeling, visual quality assessment, image retargeting, computer vision, and 3D image/video processing.

Prof. Fang serves as an Associate Editor of IEEE ACCESS and is on the Editorial Board of *Signal Processing: Image Communnication*.

Shiqi Wang received the B.S. degree in computer science from Harbin Institute of Technology, Harbin, China, in 2008, and the Ph.D. degree in computer application technology from the Peking University, Beijing, China, in 2014.

He was previously a Postdoctoral Fellow with the Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON, Canada. He is currently a Research Fellow with the Rapid-Rich Object Search Laboratory, Nanyang Technological University, Singapore. He has authored or coauthored more than 80 technical articles in refereed journals and proceedings in the areas of image and image/video coding, processing, quality assessment, and analysis. He has also proposed more than 30 technical proposals to ISO/MPEG, ITU-T, and AVS video coding standards.